REMARKS

Claims 61 and 70-93 are pending in this Application. The Applicant has cancelled claims 62-69 without prejudice to the right to pursue the subject matter of these claims in this or other applications. Applicant has added new claims 70-93 which more clearly define the subject matter of the invention and properly fall within the subject matter of the elected claims. Support for newly added claims 70-93 is found throughout the specification, in particular in canceled claims 62-69, in originally filed claims 15 and 16, and in paragraphs [0120] to [0124] as well as paragraphs [0411] and [0412]. No new matter has been entered.

Objections

112, 2nd paragraph

Claims 63, 64, 65, 66, 67, and 68 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to point out and distinctly claim the subject matter which the Applicant regards as the invention. More particular the phrase "unfractionated samples of lysed blood" has been objected to.

Claims 63, 64, 65, 66, 67, and 68 have been cancelled by the Applicant, see above. However, the Applicant respectfully traverses the rejection as it would apply to any of the newly added and/or amended claims. Applicant notes that the many embodiments of blood samples disclosed in the specification do not render the referenced phrase indefinite. However, for the purposes of expediting prosecution, Applicant has deleted the phrase "unfractionated samples of lysed blood" from the pending claims, and replaced it with the phrase "unfractionated cells of a lysed blood sample", as noted in newly added claims 75, 76, 77, 81, 82 and 83. The phrase "unfractionated cells of a lysed blood sample" is supported, for example, by Example 5, paragraph [0228] of the published application US20040241727 (hereinafter the "Published Application"), which, as noted in the instant office action, includes a centrifugation step after lysis whereby the resulting pellet containing RNA is then further utilized for quantitative PCR.

In view of this amendment and remarks clarifying the claimed embodiments, Applicant respectively requests that this rejection be reconsidered and withdrawn.

112,1st paragraph, written description

Claim 63, 64, 65, 66, 67, and 68 are rejected under 35 U.S.C. 112, first paragraph as failing to comply with the written description requirement.

The office action states that the limitation "unfractionated samples of lysed blood" appears to be new matter. Applicant traverses the rejection, but has removed the referenced phrase from the pending claims, solely for the purposes of advancing prosecution.

In view of this amendment and remarks, Applicant respectively requests that this rejection be reconsidered and withdrawn.

112,1st paragraph, enablement

Claims 61-69 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the enablement requirement.

The Applicant respectfully traverses the rejection.

Nature of the Invention and Scope of claims

The office action states:

> "the independent claim, as written, states that a comparison of a human test subject BTG2RNA level in a blood sample to a control indicates that schizophrenia is present in the test subject"
>
> "the claims are extremely broad because they require set forth that any or all comparisons between a test subject and RNA level from "control subjects" is indicative of disease."
>
> and "control subjects would could encompass patients with schizophrenia, healthy patients, patients with some other disease, such as depression or rheumatoid arthritis or multiple

sclerosis, and set forth that the comparison alone is sufficient to indicate schizophrenia, no matter the result of the comparison."

see p. 4 and p. 5 of the office action dated March 30, 2007 (hereinafter the "Office Action")

The Applicant respectfully disagrees that any comparison is sufficient to indicate the presence of schizophrenia in the test subject particularly in light of the newly amended claims. The Applicant has amended claim 61 (and corresponding independent claims 70 and 71) so as to require that the comparison of the level of BTG2 RNA in the blood sample of the test subject, with the level in blood of control subjects having schizophrenia, results in a "statistically significant similarity" before there is considered to be an indication of schizophrenia in the test subject. Newly added claim 72 (and corresponding independent claims 73 and 74) require that there be a comparison of the level of BTG2 RNA in the test subject with the level in **both (i)** control subjects not having the schizophrenia **and** (ii) control subjects having the schizophrenia. Furthermore, the comparison must result in (i) a "statistically significant similarity" between the level of RNA in the blood sample of the test subject as compared with the level of RNA in blood of the control subjects having schizophrenia **and** (ii) a "statistically significant difference" between the level of RNA in the blood sample of the test subject, and the level of RNA in blood of the control subjects not having the schizophrenia, in order to be indicative of the schizophrenia in the test subject. Newly added claim 64 (and corresponding independent claims 67 and 70) similarly require that there be **both** a "statistically significant difference" between the level of RNA in the blood sample of the test subject and the level of RNA in blood of healthy control subjects **and** a "statistically significant similarity" between the level of RNA in the blood sample of the test subject and the level of RNA in blood of control subjects who have the schizophrenia.

Thus the "control subjects" do not necessarily encompass patients with schizophrenia, healthy patients, or patients with some other disease, such as depression, rheumatoid arthritis or multiple sclerosis, as suggested at p.5 of the Office Action. Rather the control subjects are selected to either have the schizophrenia, not have the schizophrenia or are healthy control subjects. Furthermore, the comparison alone, no

matter the result of the comparison, is not sufficient to indicate schizophrenia as suggested at p. 5 of the Office Action. Instead, the comparison of the levels of the test subject with at least one set of the defined control subjects must result in a significant similarity, and in some cases, the test subject is being compared both with a negative and a positive control and a determination of a significant similarity with the positive control (ie having schizophrenia) and a significant difference with the negative control (ie not having schizophrenia) results in the determination that is indicative of schizophrenia in said test subject. Furthermore, the similarity or difference must be one with a statistical degree of significance, as determined by the many statistical techniques widely used in assessing the use of specific biomarkers in diagnosis, including those statistical techniques referenced in the instant specification, and incorporated by reference. Therefore the methods as outlined in the independent claims do not permit "any level and direction of difference in gene expression between the tested subjects is indicative of disease" as suggested at p.5 of the Office Action.

Differential Expression

The office action states that the claims do not "set forth the direction of the difference necessary to indicate schizophrenia" (p. 5 of the Office Action) and suggests that without providing this information, the mere observation of differences is an unpredictable indicator of schizophrenia.

The Applicant respectfully submits that the invention is taught in such terms that one skilled in the art can make and use the claimed invention, including the use of the elected biomarker BTG2, as an indicator of schizophrenia, without claiming the direction or the level of difference that must exists between patients having schizophrenia and individuals not having schizophrenia. The Applicant has identified the elected gene BTG2 as differentially expressed as between individuals diagnosed as having schizophrenia and individuals not having schizophrenia by demonstrating a statistical difference in the level of RNA, as described in Example 27. The statistical significance of BTG2's differential expression is evidenced by its p value of 0.0076 as listed in Table 3Y, acknowledged by the Office Action (p. 6). The Applicant has also disclosed that the

level of expression of BTG2 RNA in patients having schizophrenia is on average 2.46 fold greater than the level of expression of BTG2 RNA in healthy control patients, as further acknowledged by the Office Action (p.6). Therefore the Applicant has taught that there is a significant difference in differential expression for BTG2 as between a population of individuals having schizophrenia and a population of healthy individuals, and further has taught to compare the level of expression of BTG2 in a test individual with populations having schizophrenia, and populations not having schizophrenia using classification methods to determine the similarity or difference in gene expression levels as between the test subject and the tested populations (see paragraphs [0134] to [0137]; [0410] to [0412] and [0559] to [0561]. All of the claims require that the level of expression of RNA corresponding to BTG2 be compared with the level of BTG2 in other individuals who have schizophrenia and require, at minimum, a statistically significant similarity as between the test subject and control subjects having schizophrenia before the level of gene expression of BTG2 is considered to be indicative of schizophrenia.

Furthermore, the Applicant contends that the difference in direction and/or the fold change difference does not need to be included as a limitation within the claims to enable the claimed invention. It does not require undue experimentation for one of skill to measure a population of individuals having schizophrenia and population of individuals not having schizophrenia and determine what constitutes a statistically significant difference and a statistically significant similarity by following the methods as outlined in the Published Application. Given the widely established and validated analytical tools for analyzing gene expression levels, and the reduction to practice of the similar experiment within the Published Application, this type of experimentation qualifies as routine experimentation and therefore is not undue ('The test is not merely quantitative, since a considerable amount of experimentation is permissible, if it is merely routine, or if the specification in question provides a reasonable amount of guidance with respect to the direction in which the experimentation should proceed.'" In re Wands, 858 F.2d 731, 737, 8 USPQ2d 1400, 1404 (Fed. Cir. 1988) (citing In re Angstadt, 537 F.2d 489, 502-04, 190 USPQ 214, 217-19 (CCPA 1976)).

The Office Action also states that claims which include control subjects who do not have schizophrenia are problematic because the control group can be inclusive of individuals who have manic depression syndrome, and for this embodiment, "the specification does not provide information about an essential aspect of the invention, namely, the nature of the difference in expression that was observed between schizophrenia patients and manic depression syndrome patients".

The fact that Applicant discloses that the BTG2 gene is also differentially expressed as between individuals having schizophrenia and individuals having manic depression syndrome is not detrimental to either the value or enablement of the use of BTG2 gene as a biomarker which is indicative of schizophrenia. In accordance with claims 71, 73 and 76, the comparison of a test subject having schizophrenia with control subjects not having schizophrenia (even if this control population is solely made up of individuals having manic depression) will demonstrate a statistically significant difference and a statistically significant similarity with control subjects having schizophrenia. The comparison of a test subject who has manic depression syndrome, however, will not demonstrate a statistically significant difference as compared with control subjects not having schizophrenia (if this control population is made up solely of individuals having manic depression) and will also not demonstrate a statistically significant similarity with control subjects having schizophrenia. Thus, the requirement within the claims of a statistically significant similarity as between the test subject and control subjects having schizophrenia ensures that the level of expression being detected is selectively indicative of schizophrenia and not any other disease condition. The fact that the Applicant has demonstrated that BTG2 is differentially expressed in a statistically significant manner as between individuals having schizophrenia and individuals having manic depression syndrome as is evidenced by the p value of 0.0013 is merely further evidence that the level and/or direction of expression in patients with schizophrenia would not be statistically similar to the level of expression in patients with other non-related diseases.

Specificity of the Elected Biomarker

The Office Action states that the "specification does not establish that any particular level of expression of BTG2 (relative level or raw level) is sufficient to DETECT schizophrenia **to the exclusion of other disorders,** which is encompassed by the instant claims" (p.7-8 of the Office Action). The Office Action cites Dangond et al. (US2004/0018522) as demonstrating that "BTG2 is differentially upregulated in blood of patients with MS versus a group of controls that include healthy patients and patients with ALS (amyotrophic lateral sclerosis)" (p.8 of the Office Action) and further cites Pittman et al. (US2003/0154032) as demonstrating that BTG2 is upregulated 2.15 fold in the blood of patients with rheumatoid arthritis versus healthy controls" (p.8 of the Office Action). Thus the Examiner suggests that in order for the claims to be enabled, BTG2 must be a biomarker which is unique to schizophrenia to the exclusion of all other biomarkers.

The Applicant respectfully disagrees that the claims require that BTG2 be sufficient to detect schizophrenia to the exclusion of all other disorders. The use of a biomarker as an indication of disease, is typically just one aspect of a multi-factorial process used for diagnosing the patient with the disease. For example, as noted in Stedman's 27th Edition Medical Dictionary, "indication" is not equated with "diagnosis". The term *"indication"* is understood to mean "**the basis for initiation** of a treatment for a disease or **of a diagnostic test"** (p. 892). Even a *"diagnostic test"* is not considered to result in an absolute certainty of a diagnosis – but rather is noted as "**relating to or aiding in diagnosis**". As noted in Harrison's Principles of Internal Medicine, Introduction to Clinical Medicine "the purpose of performing a test on a patient is to reduce uncertainty about the patient's diagnosis or prognosis and to aid the clinician in making management decisions" (Ch I, pg. 11). This same text further notes that while "a perfect test would have a sensitivity of 100% and a specificity of 100% and would completely separate patients with disease from those without it...there are no perfect tests, after every test is completed the true disease state of the patient remains uncertain" (Ch I, pg. 11). Therefore, the possibility that a person with, for example, rheumatoid arthritis might be mischaracterized as having schizophrenia, although highly unlikely as

described below, does not detract from the utility of the biomarkers as an indication of schizophrenia. Rather such a hypothetical result would merely reduce the specificity of the biomarker, without affecting the sensitivity.

The amended claims all require that the level of expression of the test subject be compared with the level of expression of individuals having schizophrenia and the comparison result in a statistically significant similarity to be indicative of schizophrenia. Therefore, in order for the test to incorrectly indicate the presence of schizophrenia in a test individual, there must be another disease state which results in a statistically significant similarity in the level of expression as compared to schizophrenia. Neither Pittman et al., nor Dangond et al. demonstrate that there is a statistically significant similarity as between levels of expression of BTG1 in subjects having schizophrenia as compared with subjects having rheumatoid arthritis, amyotrophic lateral sclerosis, or multiple sclerosis.

With respect to Pittman et al., although Pittman et al. suggests a 2.1 fold increase in expression of BTG1 RNA in blood samples of individuals having rheumatoid arthritis as compared with normal individuals, the Applicant notes that the blood samples used by Pittman et al. are fractionated blood samples, such that RNA is isolated only from PBMCs (peripheral blood mononuclear cells) (see paragraph [0319] to [0320] in US2003/0154032). This is in contrast to the experiments performed by the Applicant which utilize RNA from blood samples which include all RNA expressing cells (paragraph [0410] of the Published Application). As noted in post-filing reference Du et al., each of the blood cell types can provide its own unique contribution towards a measured level of expression as between disease and control subjects ("several blood RNA isolation methods have been used to date...however, the RNA isolated using these methods comes from various blood cell subsets that originate from different developmental lineages, perform separate and distinct biological functions, and, most likely, have very different genomic expression signatures" (see p.701, 1st column of Du et al.)), therefore it is not known what the level of difference in expression would be in whole blood when comparing rheumatoid arthritis to control individuals.

Dangond et al. does not indicate the fold change level of differential expression as between patients having MLS and patients who are normal, or who have ALS, however Dangond notes that there is a difference of expression as between even the closely related diseases of MLS and ALS (see Table 8, page 27 of US2004/0018522, left column, item 6) suggesting that the expression levels of BTG2 are disease specific. This is further supported by the Applicant's own experiments which note a difference of expression for BTG2 as between schizophrenia and manic depression syndrome.

Based on the above, the Applicant has demonstrated that it is highly unlikely that the level of expression of BTG2 will be statistically similar between patients having schizophrenia and patients having rheumatoid arthritis or multiple sclerosis. Irrespective, however, the Applicant notes that even if rheumatoid arthritis or multiple sclerosis patients do demonstrate a statistically significant similarity in levels of RNA as compared with schizophrenia patients, at worst this would merely result in the BTG2 biomarker having a reduced specificity, which does not affect either the utility or the enablement of the marker.

Predictability

The Office Action suggests that "observing differences in expression between two populations is a highly unpredictable endeavor" (p.7 of the Office Action) and cites Iwamoto et al. as teaching that "expression profiling in psychiatric fields have been notoriously discordant" (p.8 of Office Action); Tsuang et al. as cautioning that "results must be interpreted with caution given several limitations including small sample size" (p.8 of the Office Action) and Vawter et al as teaching "genes that are significant by a t-test may not exceed the threshold for fold of change to be considered above background expression" (p.9 of the Office Action).

Applicant respectfully indicates that neither Iwamoto et al., nor Tsuang et al., refer to any studies which contradict the relevant teachings of the specification. Vawter et al., similarly, does not provide any teachings relating to BTG1 expression. Secondly, with respect to use of blood samples, as relates to the instant claims, Iwamoto et al. merely addresses the issue of discordance in experimental results which are due to

technical differences in the way experiments are performed between different studies performed by different researchers. Iwamoto et al. does not suggest that experimental results obtained from blood samples, including those described in the specification and in Tsuang et al. are invalid. Iwamoto et al., in fact, teach that in contrast to the blood-based instant claims, studies performed using brain tissue are particularly prone to inconsistencies due to additional factors specific to post-mortem brain tissue sampling, such as poor quality RNA and high anatomical/cellular heterogeneity of samples (e.g. Table 1). In the case of Vawter et al., Applicant wishes to point out that the cited passages relating to data inconsistencies in fact refer to pre-Vawter et al. studies which were performed using samples of pre-frontal cortex and which suffer from the drawbacks of using post-mortem brain samples which, similarly to those described by Iwamoto et al., include high variability in mRNA integrity and anatomical/cellular heterogeneity of samples. Vawter et al. at p. 42 in fact teaches that use of blood samples overcomes the inconsistencies particular to use of post-mortem brain tissue samples, and thereby in fact supports Applicant's position that the instant claims are enabled.

Applicant respectfully submits that the ordinarily skilled artisan will know to follow the comprehensive technical guidance provided in the specification (refer, for example, to the Examples section of the specification) so as to avoid the discordance issues raised by Iwamoto et al. Applicant additionally points out that Iwamoto et al. concludes at the last paragraph with the following statement, strongly supporting Applicant's position that the instantly claimed methods should be presumed enabled: "Despite the fact that multiple confounding factors complicate the findings in gene expression profiling in the clinical samples, it is one of the strongest methodologies to reveal the molecular basis of mental disorders, and its importance cannot he overemphasized."

With respect to Tsuang et al., the Examiner states that this reference cautions that the results set forth in the specification must be interpreted with caution due to various potential limitations. Applicant respectfully submits, however, that the preponderance of the teachings of Tsuang et al. are nevertheless clearly in favor of experimental data similar to that disclosed in the current Application, as being reliable. In particular, Tsuang et al. clearly teaches that the results are most likely reliable despite the limitations

cited by the Examiner, in accordance with the citation: "Despite these limitations, this work demonstrates the potential utility of blood-based RNA profiling as a diagnostic tool..." (concluding paragraph of Tsuang et al.). Applicant further submits that the experimental results disclosed in Tsuang et al. should enjoy a strong presumption of validity in view of this reference being a high-level and peer-reviewed academic publication. Applicant wishes to point out that the cautionary statements set forth in Tsuang et al. which were cited by the Examiner clearly represent a maximally conservative interpretation of the data, in line with the maximally conservative standards, for example, of the U.S. FDA. The Applicant respectfully indicates that it is improper to incorporate the standards for use by the FDA for purposes of determining patentability (see for example Application of Anthony, 56 C.C.P.A. 1443, 414 F.2d 1383, 162 U.S.P.Q. (BNA) 594 (1969); "We believe that Congress has recognized this problem and has clearly expressed its intent to give statutory authority and responsibility in this area to Federal agencies different than that given to the Patent Office. This is so because the standards established by statute for the advertisement, use, sale or distribution of drugs are quite different than the requirements under the Patent Act for the issuance of a patent."

With regards to the teachings of Vawter et al., the Applicant respectfully submits that Vawter et al. does not teach that significance as determined by t-test and fold change must both be considered for the gene to be differentially expressed. Namely, Vawter et al. at page 44, second column, teaches assignment of three independent scores for each gene depending on gene expression levels of samples and controls: classification of expression levels for both samples and controls as being above-background or not; fold-change in expression levels between samples and controls; and the p-value (t-test) for differential expression between samples and controls. Vawter et al. then classifies genes that satisfy the following arbitrary criteria as being differentially expressed between sample and control: classification of expression levels of both samples and controls as being above-background; $p<0.05$; and fold-change in expression levels between samples and controls of at least 1.4-fold. As such, the passage cited by the Examiner merely indicates that among the set of genes differentially expressed with $p<0.05$ were 21 genes whose fold-changes in expression were below 1.4-fold and whose expression levels were

below background. The Applicant submits that the fold-change criterion according to which Vawter et al. teaches that a gene is differentially expressed is purely arbitrary and should not be interpreted as a teaching that p value is not sufficient to demonstrate differential expression. Even so, however, the fold-change exhibited by BTG2 does exceed this arbitrary value. The Applicant also submits that elimination of raw expression level data which does not exceed background is a sufficiently elementary and basic procedure as to be routinely omitted from descriptions of gene expression analysis, as is the case with the current specification. Applicant considers that the discussion by Vawter et al. of data for genes whose expression levels in samples and controls are below background levels is highly unusual in the art, and does provide any meaningful teaching, as it is standard and basic procedure in the art to filter out such artifacts.

As stated in the Manual of Patent Examining Procedure at 2164.03: the "predictability or lack thereof" in the art refers to the ability of one skilled in the art to extrapolate the disclosed or known results to the claimed invention. In this case the disclosed result is a statistically significant difference in the level of BTG2 RNA as between individuals having schizophrenia and healthy individuals and a statistically significant difference in the level of BTG2 RNA as between individuals having schizophrenia and individuals having manic depression syndrome. The claimed invention requires a statistically significant similarity between the level of expression of BTG2 between the test subject and individuals having schizophrenia so as to be indicative of schizophrenia in the test subject. One skilled in the art can readily anticipate that there is similarity as between the level of expression in the test subject and a level of expression in patients having schizophrenia – knowing that the level is significantly different between subjects having schizophrenia and subjects not having schizophrenia, then there is predictability in the art.

In light of the amendments and above remarks, the Applicant contends that the claims are fully enabled, and respectfully request reconsideration and withdrawal of the instant rejection.

Conclusion

Applicant submits that all claims are allowable as written and respectfully request early favorable action by the Examiner. No new matter is added. If the Examiner believes that a telephone conversation with Applicant's attorney/agent would expedite prosecution of this application, the Examiner is cordially invited to call the undersigned attorney/agent of record.

Date: July 2, 2007

Respectfully submitted, Amy DeCloux

Amy DeClouse 54649 for

Name: Kathleen Williams
Registration No.: 34, 380
Customer No.: 29933
Edward Angell Palmer & Dodge LLP
P.O. Box 55874
Boston, MA 02205
Tel: 617-239-0100

Encl.

Excerpts from Stedman, Thomas Lathrop, 1853-1938, Stedman's Medical Dictionary 27th Edition, ed. Lippincott Williams & Wilkins, p. 492, 892,
Excerpts from Harrison's Principles of Internal Medicine, ch I Introduction to Clinical Medicine, p. 11.
Du et al. (2006) Genomic profiles for human peripheral blood T cells, B cells, natural killer cells, monocytes, and polymorphonuclear cells: Comparisons to ischemic stroke, migraine and Tourette syndrome) Genomics 87 693-703)

Application No. (if known): 10/812,731 Attorney Docket No.: 2055P(204231)

Certificate of Express Mailing Under 37 CFR 1.10

I hereby certify that this correspondence is being deposited with the United States Postal Service as Express Mail, Airbill No. EV756459975US in an envelope addressed to:

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

on July 2, 2007
Date

Amy DeCloux
Signature

Amy DeCloux
Typed or printed name of person signing Certificate

54,849
Registration Number, if applicable

(617) 439-4444
Telephone Number

Note: Each paper must have its own certificate of mailing, or this certificate must identify each submitted paper.

Amendment/Reply
3 references
Return Receipt Postcard
Amendment Transmittal
Transmittal (1 page)

EDITORS OF PREVIOUS EDITIONS

T. R. HARRISON
Editor-in-Chief, Editions 1, 2, 3, 4, 5

W. R. RESNICK
Editor, Editions 1, 2, 3, 4, 5

M. M. WINTROBE
Editor, Editions 1, 2, 3, 4, 5
Editor-in-Chief, Editions 6, 7

G. W. THORN
Editor, Editions 1, 2, 3, 4, 5, 6, 7
Editor-in-Chief, Edition 8

R. D. ADAMS
Editor, Editions 2, 3, 4, 5, 6, 7, 8, 9, 10

P. B. BEESON
Editor, Editions 1, 2

I. L. BENNETT, JR.
Editor, Editions 3, 4, 5, 6

E. BRAUNWALD
Editor, Editions 6, 7, 8, 9, 10, 12, 13, 14
Editor-in-Chief, Edition 11

K. J. ISSELBACHER
Editor, Editions 6, 7, 8, 10, 11, 12, 14
Editor-in-Chief, Editions 9, 13

R. G. PETERSDORF
Editor, Editions 6, 7, 8, 9, 11, 12, 13
Editor-in-Chief, Edition 10

J. D. WILSON
Editor, Editions 9, 10, 11, 13, 14
Editor-in-Chief, Edition 12

J. B. MARTIN
Editor, Editions 10, 11, 12, 13, 14

A. S. FAUCI
Editor, Editions 11, 12, 13
Editor-in-Chief, Edition 14

R. ROOT
Editor, Edition 12

D. L. KASPER
Editor, Editions 13, 14

S. L. HAUSER
Editor, Edition 14

D. L. LONGO
Editor, Edition 14

HARRISON'S 15TH EDITION

PRINCIPLES OF INTERNAL MEDICINE

EDITORS

EUGENE BRAUNWALD, MD, MD(Hon), ScD(Hon)
Distinguished Hersey Professor of Medicine, Faculty Dean for Academic Programs at Brigham and Women's Hospital and Massachusetts General Hospital, Harvard Medical School; Vice-President for Academic Programs, Partners HealthCare Systems, Boston

ANTHONY S. FAUCI, MD, ScD(Hon)
Chief, Laboratory of Immunoregulation; Director, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda

DENNIS L. KASPER, MD, MA(Hon)
William Ellery Channing Professor of Medicine, Professor of Microbiology and Molecular Genetics, Executive Dean for Academic Programs, Harvard Medical School; Director, Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, Boston

STEPHEN L. HAUSER, MD
Betty Anker Fife Professor and Chairman, Department of Neurology, University of California San Francisco, San Francisco

DAN L. LONGO, MD
Scientific Director, National Institute on Aging, National Institutes of Health, Bethesda and Baltimore

J. LARRY JAMESON, MD, PHD
Irving S. Cutter Professor and Chairman, Department of Medicine, Northwestern University Medical School; Physician-in-Chief, Northwestern Memorial Hospital, Chicago

McGraw-Hill
MEDICAL PUBLISHING DIVISION

New York San Francisco Washington, DC Auckland Bogotá Caracas Lisbon London
Madrid Mexico City Milan Montreal New Delhi San Juan Singapore Sydney Tokyo Toronto

McGraw-Hill
A Division of The McGraw-Hill Companies

Note: Dr. Fauci and Dr. Longo's works as editors and authors were performed outside the scope of their employment as U.S. government employees. These works represent their personal and professional views and not necessarily those of the U.S. government.

HARRISON'S PRINCIPLES OF INTERNAL MEDICINE

Harrison's
PRINCIPLES OF INTERNAL MEDICINE
Fifteenth Edition

1234567890 DOWDOW 0987654321
ISBN 0-07-007272-8 (Combo)
0-07-007273-6 (Vol. 1)
0-07-007274-4 (Vol. 2)
0-07-913686-9 (Set)

FOREIGN LANGUAGE EDITIONS
Arabic (Thirteenth Edition)—McGraw-Hill Libri Italia srl (est. 1996)
Chinese (Twelfth Edition)—McGraw-Hill Book Company—Singapore © 1994
Croatian (Thirteenth Edition)—Placebo, Split, Croatia
French (Fourteenth Edition)—McGraw-Hill Publishing Co., Maidenhead, UK © 1999
German (Fourteenth Edition)—McGraw-Hill Publishing Co., Maidenhead, UK © 1999
Greek (Fourteenth Edition)—Parissianos, Athens, Greece © 2000
Italian (Fourteenth Edition)—McGraw-Hill Libri Italia srl, Milan © 1999
Japanese (Eleventh Edition)—Hirokawa © 1991
Polish (Fourteenth Edition)—Czelej Publishing Company, Lubin, Poland (est. 2000)
Portuguese (Fourteenth Edition)—McGraw-Hill Interamericana do Brasil Ltda © 1998
Romania (Fourteenth Edition)—Teora Publishers, Bucharest, Romania (est. 2000)
Spanish (Fourteenth Edition)—McGraw-Hill Interamericana de Espana, Madrid © 1998
Turkish (Thirteenth Edition)—McGraw-Hill Libri Italia srl (est. 1996)

This book was set in Times Roman by Progressive Information Technologies. The editors were Martin Wonsiewicz and Mariapaz Ramos Englis. The production director was Robert Laffler. The index was prepared by Irving C. Tullar. The text and cover designer was Marsha Cohen/Parallelogram Graphics.

R. R. Donnelley and Sons, Inc. was the printer and binder.

Library of Congress Cataloging-in-Publication Data

Harrison's principles of internal medicine—15th ed./editors, Eugene Braunwald . . . [et al.]. p. cm.
Includes bibliographical references and index.
ISBN 0-07-913686-9 (set)—ISBN 0-07-007273-6 (v. 1)—ISBN 0-07-0072744-4 (v. 2)
1. Internal medicine. I. Braunwald, Eugene, date
RC46.H333 2001
616—dc21

00-063809

reduces the likelihood of hyperthyroidism in a patient with paroxysmal atrial fibrillation.

While the representativeness and availability heuristics may play the major roles in shaping early diagnostic hypotheses, the acuity of a patient's illness can also be very influential. For example, clinicians are taught to consider aortic dissection routinely as a possible cause of acute severe chest discomfort along with myocardial infarction, even though the typical history of dissection is different from myocardial infarction and dissection is far less prevalent (Chap. 247). This recommendation is based on the recognition that a relatively rare but catastrophic diagnosis like aortic dissection is very difficult to make unless it is explicitly considered. If the clinician fails to elicit any of the characteristic features of dissection by history and finds equivalent blood pressures in both arms and no pulse deficits, he or she may feel comfortable in discarding the aortic dissection hypothesis. If, however, the chest x-ray shows a widened mediastinum, the hypothesis may be reinstated and a diagnostic test ordered [e.g., thoracic computed tomography (CT) scan, transesophageal echocardiogram] to evaluate it more fully. In noncritical situations, the prevalence of potential alternative diagnoses should play a much more prominent role in diagnostic hypothesis generation. The value of conducting a rapid systematic clinical survey of symptoms and organ systems to avoid missing important but inapparent clues cannot be overstated.

Because the generation and evaluation of appropriate diagnostic hypotheses is a skill that not all clinicians possess to an equal degree, errors in this process can occur, and in the patient with serious acute illness these may lead to tragic consequences. Consider the following hypothetical example. A 45-year-old male patient with a 3-week history of a "flulike" upper respiratory infection (URI) presented to his physician with symptoms of dyspnea and a productive cough. Based on the presenting complaint, the clinician pulled out a "URI Assessment Form" to improve quality and efficiency of care. The physician quickly completed the examination components outlined on this structured form, noting in particular the absence of fever and a clear chest examination. He then prescribed an antibiotic for presumed bronchitis, showed the patient how to breathe into a paper bag to relieve his "hyperventilation," and sent him home with the reassurance that his illness was not serious. After a sleepless night with significant dyspnea unrelieved by rebreathing into a bag, the patient developed nausea and vomiting and collapsed. He was brought into the Emergency Department in cardiac arrest and could not be resuscitated. Autopsy showed a posterior wall myocardial infarction and a fresh thrombus in an atherosclerotic right coronary artery. What went wrong? The clinician decided, even before starting the history, that the patient's complaints were not serious. He therefore felt confident that he could perform an abbreviated and focused examination using the URI assessment protocol rather than considering the full range of possibilities and performing appropriate tests to confirm or refute his initial hypotheses. In particular, by concentrating on the "URI," the clinician failed to elicit the full dyspnea history, which would have suggested a far more serious disorder, and did not even search for other symptoms that could have directed him to the correct diagnosis.

This example illustrates how patients can diverge from textbook symptoms and the potential consequences of being unable to adapt the diagnostic process to real-world challenges. The expert, while recognizing that common things occur commonly, approaches each evaluation on high alert for clues that the initial diagnosis may be wrong. Patients often provide information that "does not fit" with any of the leading diagnostic hypotheses being considered. Distinguishing real clues from false trails can only be achieved by practice and experience. A less experienced clinician who tries to be too efficient (as in the above example) can make serious judgment errors.

MAJOR INFLUENCES ON CLINICAL DECISION-MAKING More than a decade of research on variations in clinician practice patterns has shed much light on forces that shape clinical decisions. The use of heuristic "shortcuts," as detailed above, provides a partial explanation, but several other key factors play an important role in shaping diagnostic hypotheses and management decisions. These factors can be grouped conceptually into three overlapping categories: (1) factors related to physician personal characteristics and practice style, (2) factors related to the practice setting, and (3) economic incentive factors.

Practice Style Factors One of the key roles of the physician in medical care is to serve as the patient's agent to ensure that necessary care is provided at a high level of quality. Factors that influence this role include the physician's knowledge, training, and experience. It is obvious that physicians cannot practice evidence-based medicine if they are unfamiliar with the evidence. As would be expected, specialists generally know the evidence in their field better than do generalists. Surgeons may be more enthusiastic about recommending surgery than medical doctors because their belief in the beneficial effects of surgery is stronger. For the same reason, invasive cardiologists are much more likely to refer chest pain patients for diagnostic catheterization than are noninvasive cardiologists or generalists. The physician beliefs that drive these different practice styles are based on personal experience, recollection, and interpretation of the available medical evidence. For example, heart failure specialists are much more likely than generalists to achieve target angiotensin-converting enzyme (ACE) inhibitor therapy in their heart failure patients because they are more familiar with what the targets are (as defined by large clinical trials), have more familiarity with the specific drugs (including dosages and side effects), and are less likely to overreact to foreseeable problems in therapy such as a rise in creatinine levels or symptomatic hypotension. Other intriguing research has shown a wide distribution of acceptance times of antibiotic therapy for peptic ulcer disease following widespread dissemination of the "evidence" in the medical literature. Some gastroenterologists accepted this new therapy before the evidence was clear (reflecting, perhaps, an aggressive practice style), and some gastroenterologists lagged behind (a conservative practice style, associated in this case with older physicians). As a group, internists lagged several years behind gastroenterologists.

The opinion of influential leaders can also have an important effect on practice patterns. Such influence can occur at both the national level (e.g., expert physicians teaching at national meetings) and the local level (e.g., local educational programs, "curbside consultants"). Opinion leaders do not have to be physicians. When conducting rounds with clinical pharmacists, physicians are less likely to make medication errors and more likely to use target levels of evidence-based therapies.

The patient's welfare is not the only concern that drives clinical decisions. The physician's perception about the risk of a malpractice suit resulting from either an erroneous decision or a bad outcome creates a style of practice referred to as *defensive medicine*. This practice involves using tests and therapies with very small marginal returns to preclude future criticism in the event of an adverse outcome. For example, a 40-year-old woman who presents with a long-standing history of intermittent headache and a new severe headache along with a normal neurologic examination has a very low likelihood of structural intracranial pathology. Performance of a head CT or magnetic resonance imaging (MRI) scan in this situation would constitute defensive medicine. On the other hand, the results of the test could provide reassurance to an anxious patient.

Practice Setting Factors Factors in this category relate to the physical resources available to the physician's practice and the practice environment. *Physician-induced demand* is a term that refers to the repeated observation that physicians have a remarkable ability to accommodate to and employ the medical facilities available to them. A classic early study in this area showed that physicians in Boston had an almost 50% higher hospital admission rate than did physicians in New Haven, despite there being no obvious differences in the health of the cities' inhabitants. The physicians in New Haven were not aware of using fewer hospital beds for their patients, nor were the Boston physicians aware of using less stringent criteria to admit patients.

Other environmental factors that can influence decision-making include the local availability of specialists for consultations and procedures, "high tech" facilities such as angiography suites, a heart surgery program, and MRI machines.

Economic Incentives Economic incentives are closely related to the other two categories of practice-modifying factors. Financial issues can exert both stimulatory and inhibitory influences on clinical practice. In general, physicians are paid on a fee-for-service, capitation, or salary basis (Chap. 4). In fee-for-service, the more the physician does, the more the physician gets paid. The incentive in this case is to do more. When fees are reduced (discounted fee-for-service), doctors tend to increase the number of services billed for. *Capitation*, in contrast, provides a fixed payment per patient per year, encouraging physicians to take on more patients but to provide each patient with fewer services. Expensive services are more likely to be affected by this type of incentive than inexpensive preventive services. Salary compensation plans pay physicians the same regardless of the amount of clinical work performed. The incentive here is to see fewer patients. Recognizing these powerful shapers of physician behavior, managed care plans have begun to explore combinations of the three reimbursement types with the goal of improving individual physician productivity while restraining their use of expensive tests and therapies.

In summary, expert clinical decision-making can be appreciated as a complex interplay between cognitive devices used to simplify large amounts of complex information interacting with physician biases reflecting education, training, and experience, all of which are shaped by powerful, sometimes perverse, external forces. In the next section, we will review a set of statistical tools and concepts that can assist in making clinical decisions under uncertainty.

QUANTITATIVE METHODS TO AID CLINICAL DECISION-MAKING

The process of medical decision-making can be divided into two parts: (1) defining the available courses of action and estimating the likely outcomes with each, and (2) assessing the desirability of the outcomes. The former task involves integrating key information about the patient along with relevant evidence from the medical literature to create the structure of a decision problem. The remainder of this chapter will present some quantitative tools to assist the clinician in these activities. These tools can be divided into those that assist the clinician in making better outcome predictions, which are then used to make decisions, and those that support the decision process directly. While these tools are not yet used routinely in daily clinical practice, the computerization of medicine is creating the required substrate for their future widespread dissemination.

QUANTITATIVE MEDICAL PREDICTIONS **Diagnostic Testing** The purpose of performing a test on a patient is to reduce uncertainty about the patient's diagnosis or prognosis and to aid the clinician in making management decisions. Although diagnostic tests are commonly thought of as laboratory tests (e.g., measurement of serum amylase level) or procedures (e.g., colonoscopy or bronchoscopy), any technology that changes our understanding of the patient's problem qualifies as a diagnostic test. Thus, even the history and physical examination can be considered a form of diagnostic test. In clinical medicine, it is common to reduce the results of a test to a dichotomous outcome, such as positive or negative, normal or abnormal. In many cases, this simplification results in the waste of useful information. However, such simplification makes it easier to demonstrate some of the quantitative ways in which test data can be used.

To characterize the accuracy of diagnostic tests, four terms are routinely used (Table 3-1). The *true-positive rate*, i.e., the sensitivity, provides a measure of how well the test correctly identifies patients with disease. The *false-negative rate* is calculated as (1 − sensitivity). The *true-negative rate*, i.e., the specificity, reflects how well the test correctly identifies patients without disease. The *false-positive rate* is

Table 3-1 Measures of Diagnostic Test Accuracy

	Disease Status	
Test Result	**Present**	**Absent**
Positive	True-positive (*TP*)	False-positive (*FP*)
Negative	False-negative (*FN*)	True-negative (*TN*)

IDENTIFICATION OF PATIENTS WITH DISEASE

True-positive rate (sensitivity) = $TP/(TP + FN)$
False-negative rate = $FN/(TP + FN)$
True-positive rate = 1 − false-negative rate

IDENTIFICATION OF PATIENTS WITHOUT DISEASE

True-negative rate (specificity) = $TN/(TN + FP)$
False-positive rate = $FP/(TN + FP)$
True-negative rate = 1 − false-positive rate

(1 − specificity). A perfect test would have a sensitivity of 100% and a specificity of 100% and would completely separate patients with disease from those without it.

Calculating sensitivity and specificity require selection of a cut-point value for the test to separate "normal" from "diseased" subjects. As the cutpoint is moved to improve sensitivity, specificity typically falls and vice versa. This dynamic tradeoff between more accurate identification of subjects with versus those without disease is often displayed graphically as a receiver operating characteristic (ROC) curve. An ROC curve plots sensitivity (*y*-axis) versus 1 − specificity (*x*-axis). Each point on the curve represents a potential cutpoint with an associated sensitivity and specificity value. The area under the ROC curve is often used as a quantitative measure of the information content of a test. Values range from 0.5 (no diagnostic information at all, test is equivalent to flipping a coin) to 1.0 (perfect test).

In the diagnostic testing literature, ROC areas are often used to compare alternative tests. The test with the highest area (i.e., closest to 1.0) is presumed to be the most accurate. However, ROC curves are not a panacea for evaluation of diagnostic test utility. Like Bayes' theorem, they are typically focused on only one possible test parameter (e.g., ST segment response in a treadmill exercise test) to the exclusion of other potentially relevant data. In addition, ROC area comparisons do not simulate the way test information is actually used in clinical practice. Finally, biases in the underlying population used to generate the ROC curves (e.g., related to an unrepresentative test sample) can bias the ROC area and the validity of a comparison among tests.

Measures of Disease Probability and Bayes' Theorem Unfortunately, there are no perfect tests; after every test is completed the true disease state of the patient remains uncertain. Quantitating this residual uncertainty can be done with Bayes' theorem. This theorem provides a simple mathematical way to calculate the posttest probability of disease from three parameters: the pretest probability of disease, the test sensitivity, and the test specificity (Table 3-2). The pretest probability is a quantitative expression of the confidence in a diagnosis before the test is performed. In the absence of more relevant information it is usually estimated from the prevalence of the disease in the underlying population. For some common conditions, such as coronary artery disease (CAD), nomograms and statistical models have been created to generate better estimates of pretest probability from elements of the history and physical examination. The posttest probability, then, is a revised statement of the confidence in the diagnosis, taking into account both what was known before and after the test.

To understand how Bayes' theorem creates this revised confidence statement, it is useful to examine a nomogram version of Bayes' theorem that uses the same three parameters to predict the posttest probability of disease (Fig. 3-1). In this nomogram, the accuracy of the diagnostic test in question is summarized by the likelihood ratio for a

STEDMAN'S Medical Dictionary

27th Edition

Illustrated in Color

LIPPINCOTT WILLIAMS & WILKINS
A Wolters Kluwer Company
Philadelphia • Baltimore • New York • London
Buenos Aires • Hong Kong • Sydney • Tokyo

Senior Managing Editor: Maureen Barlow Pugh
Managing Editor: Barbara Werner
New Terms Editor: Thomas W. Filardo, MD
Copy Editors: Peter W. Binns, Linda G. Francis, Raymond Lukens, Bonnie Montgomery
Chief On-Line Editor: Barbara L. Ferretti
On-Line Editors: Kathryn J. Cadle, Dana Workman
Proofreaders: Peter W. Binns; David A. Bloom, MD; Alfred J. Bollet, MD; Ted Burk; Regina Lavette Davis; John A. Day, Jr., MD, FCCP; Richard Diamanti; John H. Dirckx, MD; Thomas W. Filardo, MD; Linda G. Francis; John M. Last, MD, FRACP, FRCPC; Raymond Lukens; Kate Mason, CMT; Joan Sarchese
Database Programmers: Dave Marcus, Lexi-Comp Inc., Hudson, OH
Art Director: Jonathan Dimes
Illustrations: Neil O. Hardy
Additional artwork by: Mary Anna Barratt-Dimes, Kathryn Born, Rob Duckwall, Timothy Hengst, Mikki Senkarik, Michael Schenk, Larry Ward
Graphic preparation assistance: Susan Caldwell, Jennifer Clements, Thomas Dolan, Christina Nihira
Design: Dan Pfisterer

351 West Camden Street
Baltimore, Maryland 21201–2436 USA

Database design by Lexi-Comp Inc., Hudson, OH
Printed in the United States of America by World Color, Inc.

Library of Congress Cataloging-in-Publication Data

Stedman, Thomas Lathrop, 1853-1938.
Stedman's medical dictionary.—27th ed.
p.; cm.
ISBN 0-683-40007-X (regular)—ISBN 0-683-40008-8 (deluxe)
1. Medicine—Dictionaries. I. Title: Medical dictionary. II. Title.
[DNLM: 1. Medicine—Dictionary—English. W 13 S812m 1999]
R121 .S8 1999
610'.3—dc21

99-056094

00 01 02 03 04 05
1 2 3 4 5

therapeutic i., the ratio of LD_{50} to ED_{50}, used in quantitative comparison of drugs.

thoracic i., anteroposterior diameter of the thorax times 100 divided by the transverse diameter of the thorax. SYN chest i.

tibiofemoral i., the ratio obtained by multiplying the length of the tibia by 100 and dividing by the length of the femur.

transversovertical i., SYN vertical i.

tuberculoopsonic i., the opsonic i. calculated in relation to tuberculous infection, with an actively growing culture of *Mycobacterium tuberculosis* or the strain of tubercle bacillus from the patient being used in the test.

ultraviolet i., a daily i. issued by the U.S. National Weather Service for many cities, forecasting the amount of dangerous ultraviolet light that will arrive at the earth's surface about noon the following day.

uricolytic i., the percentage of uric acid oxidized to allantoin before being secreted.

vertical i., the relation of the height to the length of the skull: (height × 100)/length. SYN height-length i., length-height i., transversovertical i.

vital i., the ratio of births to deaths within a population during a given time.

Volpe-Manhold i. (V-MI), an index for comparing the amount of dental calculus in individuals.

volume i., an indication of the relative size (e.g., volume) of erythrocytes, calculated as follows: hematocrit value, expressed as per cent of normal ÷ red blood cell count, expressed as per cent of normal = volume i.

zygomaticoauricular i., the ratio between the zygomatic and the auricular diameters of the skull or head.

---

in·di·can (in'di-kan). 1. Indoxyl β-D-glucoside from *Indigofera* species and *Polygonum tinctorium;* a source of indigo. SYN plant i. 2. 3-Indoxylsulfuric acid, a substance found (as its salts) in sweat and in variable amounts in urine; indicative, when in quantity, of protein putrefaction in the intestine (indicanuria). SYN metabolic i., uroxanthin.

metabolic i., SYN indican (2).

plant i., SYN indican (1).

in·di·can·i·dro·sis (in'di-kan-i-drō'sis). Excretion of indican in the sweat. [indican + G. *hidrōs*, sweat]

in·di·cant (in'di-kant). 1. Pointing out; indicating. 2. An indication; especially a symptom indicating the proper line of treatment. [L. *in-dico*, pres. p. *-ans* (*-ant*), to point out]

in·di·can·u·ria (in'di-kan-ū'rē-ă). An increased urinary excretion of indican, a derivative of indol formed chiefly in the intestine when protein is putrefied; indol is also formed during the putrefaction of protein in other sites.

in·di·ca·tion (in-di-kā'shŭn). The basis for initiation of a treatment for a disease or of a diagnostic test; may be furnished by a knowledge of the cause (causal i.), by the symptoms present (symptomatic i.), or by the nature of the disease (specific i.). [L. fr. *in-dico*, pp. *-atus*, to point out, fr. *dico*, to proclaim]

off label i., use of a medication for a purpose other than that approved by the FDA.

in·di·ca·tor (in'di-kā-ter, -tŏr). 1. In chemical analysis, a substance that changes color within a certain definite range of pH or oxidation potential, or in any way renders visible the completion of a chemical reaction; e.g., litmus, phenolsulfonphthalein. 2. An isotope that is used as a tracer. 3. The labeled substance whose distribution between reactants of a system is used to determine the amount of analyte present. [L. one that points out]

alizarin i., a solution consisting of 1 g sodium alizarin sulfonate dissolved in 100 mL distilled water; used as an i. for free acidity in gastric contents.

clinical i., a measure, process, or outcome used to judge a particular clinical situation and indicate whether the care delivered was appropriate.

health i., variable, susceptible to direct measurement, that reflects the state of health of persons in a community.

oxidation-reduction i., a substance that undergoes a definite color change at a specific oxidation potential. SYN redox i.

redox i., SYN oxidation-reduction i.

in·di·ces (in'di-sēz). Alternative plural of index.

In·di·el·la (in-dē-el'ă). Old name for *Madurella*.

in·dig·e·nous (in-dij'ĕ-nŭs). Native; natural to the country or region where found. [L. *indigenus*, born in fr. *indu*, within (old form of *in*), + G. *-gen*, producing]

in·di·ges·tion (in-di-jes'chŭn). Nonspecific term for a variety of symptoms resulting from a failure of proper digestion and absorption of food in the alimentary tract.

acid i., i. resulting from hyperchlorhydria; often used by the laity as a synonym for pyrosis.

fat i., SYN steatorrhea.

gastric i., SYN dyspepsia.

nervous i., i. caused by emotional upsets or stress.

in·di·go (in'dĭ-gō) [C.I. 73000]. A blue dyestuff obtained from *Indigofera tinctoria*, and other species of *Indigofera* (family Leguminosae); also made synthetically. SYN indigo blue, indigotin. [L. *indicum*, fr. G. *indikon*, indigo, ntr. of *Indikos*, Indian]

in·di·go blue. SYN indigo.

in·di·go car·mine [C.I. 73015]. A blue dye used for measurement of kidney function and as a special stain for Negri bodies. SYN sodium indigotin disulfonate.

in·dig·o·tin (in-dig'ō-tin, in-di-gō'tin). SYN indigo.

in·di·go·u·ria, in·di·gu·ria (in'dĭ-gō-ū'rē-ă, in-di-goo'rē-ă). The excretion of indigo in the urine.

in·dis·po·si·tion (in-dis-pō-zish'ŭn). Illness, usually slight; malaise. [L. *in* neg. + *dispositio*, an arrangement, fr. *dis-pono*, pp. *-positus*, to place apart]

in·di·um (In) (in'dē-ŭm). A metallic element, atomic no. 49, atomic wt. 114.82. [*indigo*, because of its blue line in the spectrum]

in·di·um-111 (^{111}In). A cyclotron-produced radionuclide with a half-life of 2.8049 days and with gamma ray emissions of 171.2 and 245.3 kiloelectron volts. In a chloride form, it is used as a bone marrow and tumor-localizing tracer; in a chelate form, as a cerebrospinal fluid tracer. It is also used as a white blood cell labeling agent and as an antibody label.

i. chloride, i. trichloride, $InCl_3$; used in electron microscopy to stain nucleic acids in thin tissue sections.

in·di·um-113m (^{113m}In). A radioactive isomer of ^{113}In; it has a half-life of 1.658 hours; it has been used in cisternography and as a diagnostic aid in cardiac output.

in·di·vid·u·a·tion (in'di-vid-ū-ā'shŭn). 1. Development of the individual from the specific. 2. In jungian psychology, the process by which one's personality is differentiated, developed, and expressed. 3. Regional activity in an embryo as a response to an organizer.

in·do·cy·a·nine green (in-dō-sī'ă-nēn). A tricarbocyanine dye that binds to serum albumin and is used in blood volume determinations and in liver function tests.

in·do·cy·bin (in-dō-sī'bin). SYN psilocybin.

in·dol·ac·e·tu·ria (in'dōl-as-ĕ-too'rē-ă). Excretion of an appreciable amount of indoleacetic acid in the urine; a manifestation of Hartnup disease, also seen in patients with carcinoid tumors.

in·dol·a·mine (in-dol'ă-mēn). General term for an indole or indole derivative containing a primary, secondary, or tertiary amine group (e.g., serotonin).

in·dole (in'dōl). 1. 2,3-Benzopyrrole; basis of many biologically active substances (e.g., serotonin, tryptophan); formed in degradation of tryptophan. SYN ketole. 2. Any of many alkaloids containing the i. (1) structure.

in·do·lent (in'dō-lent). Inactive; sluggish; painless or nearly so; said of a morbid process. [L. *in-* neg. + *doleo*, pr. p. *dolens* (*-ent-*), to feel pain]

in·dol·ic ac·ids (in-dōl'ik). Metabolites of L-tryptophan formed within the body or by intestinal microorganisms; the principal ones encountered in urine are indoleacetic acid, indoleacetylglutamine, 5-hydroxyindoleacetic acid, and indolelactic acid.

antenatal d., SYN prenatal d.

clinical d., a d. made from a study of the signs and symptoms of a disease.

differential d., the determination of which of two or more diseases with similar symptoms is the one from which the patient is suffering, by a systematic comparison and contrasting of the clinical findings. SYN differentiation (2).

d. by exclusion, a d. made by excluding those diseases to which only some of the patient's symptoms might belong, leaving one disease as the most likely d., although no definitive tests or findings establish that d.

laboratory d., a d. made by a chemical, microscopic, microbiologic, immunologic, or pathologic study of secretions, discharges, blood, or tissue.

neonatal d., systematic evaluation of the newborn for evidence of disease or malformations, and the conclusion reached.

pathologic d., a d., sometimes postmortem, made from an anatomic and/or histologic study of the lesions present.

physical d., (1) a d. made by means of physical examination of the patient. (2) the process of a physical examination.

prenatal d., d. utilizing procedures available for the recognition of diseases and malformations *in utero*, and the conclusion reached. SYN antenatal d.

di·ag·nos·tic (dī-ag-nos'tik). 1. Relating to or aiding in diagnosis. 2. Establishing or confirming a diagnosis.

di·ag·nos·ti·cian (dī'ag-nos-tish'ăn). One who is skilled in making diagnoses; formerly, a name for specialists in internal medicine.

***Diagnostic and Statistical Manual of Mental Disorders* (DSM).** A system of classification, published by the American Psychiatric Association, that divides recognized mental disorders into clearly defined categories based on sets of objective criteria. Representing a majority view (rather than a consensus) of hundreds of contributors and consultants, DSM is widely recognized as a diagnostic standard and widely used for reporting, coding, and statistical purposes.

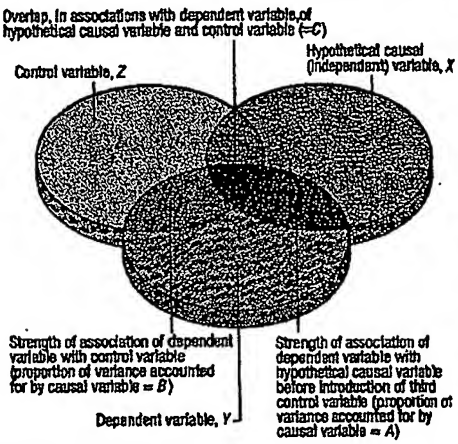
The first edition (1952), based on the sixth revision of the *International Classification of Diseases* (*ICD-6*), was intended to promote uniformity in the naming and reporting of psychiatric disorders. It contained definitions of all named disorders, but no sets of diagnostic criteria. While its classification of mental disorders showed the influence of Freudian psychoanalysis, its nomenclature (e.g., depressive reaction, anxiety reaction, schizophrenic reaction) reflected the theories of Adolf Meyer (1866–1950). The second edition (*DSM-II*, 1968) preserved the psychoanalytic orientation but dropped the "reaction" terminology. The third edition (*DSM-III*, 1980) abandoned much of the rigidly psychodynamic thinking of the earlier editions and, for the first time, provided explicit diagnostic criteria and introduced a multiaxial system whereby different aspects of a patient's condition could be separately assessed. Briefly stated, the axes are I, clinical disorders; II, personality disorders and mental retardation; III, general medical disorders; IV, psychosocial and environmental stressors; and V, overall level of functioning. A revised version of the third edition (*DSM-IIIR*, 1987) incorporated a number of improvements and clarifications. The fourth edition (*DSM-IV*) appeared in May, 1994. It follows its two predecessors closely in general outline, and like them is coordinated with and partly derived from *ICD-9*. For many observers, the most significant change in *DSM-IV* is the renaming of the category formerly called "Organic Mental Syndromes and Disorders" as "Delirium, Dementia, and Amnestic and Other Cognitive Disorders," a shift in terminology intended to avoid the implication that mental disorders in other categories are not organic.

di·a·gram. A simple, graphic depiction of an idea or object.

Dieuaide d., SYN triaxial reference *system*.

flow d., a d. composed of blocks connected by arrows representing steps in a process such as decision analysis.

Venn d., pictorial representation of the extent to which two or more quantities or concepts are mutually inclusive and exclusive.



Venn diagram

di·a·ki·ne·sis (dī'ă-ki-nē'sis). Final stage of prophase in meiosis I, in which the chiasmata present during the diplotene stage disappear, the chromosomes continue to shorten, and the nucleolus and nuclear membrane disappear. [G. *dia*, through, + *kinēsis*, movement]

dial (dī'ăl, dīl). A clock face or instrument resembling a clock face. [L. *dies*, day]

astigmatic d., a diagram of radiating lines, used to test for astigmatism.

Di·a·lis·ter (dī-ăl-is'ter). An obsolete name for a genus of bacteria, the type species of which, *D. pneumosintes*, is now placed in the genus *Bacteroides*.

di·al·lyl (dī-al'il). A compound containing two allyl groups.

di·al·y·sance (dī-al'i-sans). The number of milliliters of blood completely cleared of any substance by an artificial kidney or by peritoneal dialysis in a unit of time; conventional clearance formulas are expressed as mm/min. [fr. dialysis]

di·al·y·sate (dī-al'i-sāt). That part of a mixture that passes through a dialyzing membrane; the material that does not pass through is referred to as the retentate. SYN diffusate.

di·al·y·sis (dī-al'i-sis). 1. A form of filtration to separate crystalloid from colloid substances (or smaller molecules from larger ones) in a solution by interposing a semipermeable membrane between the solution and dialyzing fluid; the crystalloid (smaller) substances pass through the membrane into the dialyzing fluid on the other side, the colloids do not. 2. The separation of substances across a semipermeable membrane on the basis of particle size and/or concentration gradients. 3. A method of artificial kidney function. [G. a separation, fr. *dialyo*, to separate]

continuous ambulatory peritoneal d. (CAPD), method of peritoneal d. performed in ambulatory patients with influx and efflux of dialysate during normal activities.

equilibrium d., in immunology, a method for determination of association constants for hapten-antibody reactions in a system in which the hapten (dialyzable) and antibody (nondialyzable) solutions are separated by semipermeable membranes. Since at equilibrium the quantity of free hapten will be the same in the two compartments, quantitative determinations can be made of hapten-bound antibody, free antibody, and free hapten.

extracorporeal d., hemodialysis performed through an apparatus outside the body.

peritoneal d., removal from the body of soluble substances and

Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Genomics 87 (2006) 693–703
GENOMICS
www.elsevier.com/locate/ygeno

Genomic profiles for human peripheral blood T cells, B cells, natural killer cells, monocytes, and polymorphonuclear cells: Comparisons to ischemic stroke, migraine, and Tourette syndrome

XinLi Du [a,b], Yang Tang [a,b], Huichun Xu [a,b], Lisa Lit [a,b], Wynn Walker [a,b], Paul Ashwood [a,d], Jeffrey P. Gregg [a,c], Frank R. Sharp [a,b,*]

[a] *MIND Institute, University of California at Davis, 2805 50th Street, Sacramento, CA 95817, USA*
[b] *Department of Neurology, University of California at Davis, 2805 50th Street, Sacramento, CA 95817, USA*
[c] *Department of Pathology, University of California at Davis, 2805 50th Street, Sacramento, CA 95817, USA*
[d] *Department of Medicine, University of California at Davis, 2805 50th Street, Sacramento, CA 95817, USA*

Received 3 November 2005; accepted 5 February 2006
Available online 20 March 2006

Abstract

Blood genomic profiling has been applied to disorders of the blood and various organ systems including brain to elucidate disease mechanisms and identify surrogate disease markers. Since most studies have not examined specific cell types, we performed a preliminary genomic survey of major blood cell types from normal individuals using microarrays. $CD4^+$ T cells, $CD8^+$ T cells, $CD19^+$ B cells, $CD56^+$ natural killer cells, and $CD14^+$ monocytes were negatively selected using the RosetteSep antibody cocktail, while polymorphonuclear leukocytes were separated with density gradient media. Genes differentially expressed by each cell type were identified. To demonstrate the potential use of such cell subtype-specific genomic expression data, a number of the major genes previously reported to be regulated in ischemic stroke, migraine, and Tourette syndrome are shown to be associated with distinct cell populations in blood. These specific gene expression, cell-type-related profiles will need to be confirmed in larger data sets and could be used to study these and many other neurological diseases.

Keywords: Blood; Humans; Gene expression; Microarrays; Genome; T cells; B cells; NK cells; Neutrophils; Migraine; Stroke; Tourette

Gene expression profiling of peripheral blood using microarrays has been applied to malignant and immune disorders, including leukemia, lymphoma, systemic lupus erythematosis, rheumatoid arthritis, and many others [1–4]. This approach has helped identify important diagnostic and prognostic markers as well as potential therapeutic targets. This approach has also been extended to many diseases of other organ systems. It is likely that many inflammatory, autoimmune, and genetic factors could affect gene expression of peripheral blood cells without causing overt changes to hematological and immunological phenotypes. Proof-of-principle blood genomic studies have been performed in animals [5] and humans [6,7]. Subsequent studies have demonstrated characteristic blood genomic patterns for acute ischemic stroke [8], migraine headache [9], Tourette syndrome [10], renal cell carcinoma [11], multiple sclerosis [12], benzene exposure [13], trauma [14], and neurogenetic disorders including neurofibromatosis type I, tuberous sclerosis type II, Down syndrome [7,15], and Huntington chorea [16]. The study of blood gene expression profiles appears to be a promising approach that may provide mechanistic insights and surrogate markers for many diseases.

Several blood RNA isolation methods have been used to date. These include methods starting with whole blood, mononuclear cells, and buffy coat [6,7,11,14,17,18]. However, the RNA isolated using these methods comes from various blood cell subsets that originate from different

* Corresponding author. MIND Institute, University of California at Davis, 2805 50th Street, Sacramento, CA 95817, USA. Fax: +1 916 703 0369.
E-mail address: frank.sharp@ucdmc.ucdavis.edu (F.R. Sharp).

0888-7543/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2006.02.003

developmental lineages, perform separate and distinct biological functions, and, most likely, have very different genomic expression signatures. It has been recognized that age and gender and the different composition of blood cells from each individual represent a major source of normal variation of blood gene expression [6,7]. In addition, a disease may predominantly affect one specific blood cell subtype while sparing others. Therefore, characterizing the contribution of every blood cell subtype to the overall blood genomic pattern may be essential to distinguish significant genomic changes from noise, interpret the disease-related patterns, and decide on the proper blood cell types to perform follow-up confirmatory analyses.

Expression profiles of blood cells such as T lymphocytes [14,19] and platelets [18,20] have been described. However, studies that compare directly the whole genomic expression profiles of several major blood cell subtypes have not been performed in detail. In this study, we attempted to build a preliminary gene expression database by comparing major leukocyte subsets from three healthy donors, including polymorphonuclear cells (PMN), monocytes, B cells, $CD4^+$ T cells, $CD8^+$ cytotoxic T cells, and natural killer (NK) cells to determine whether there is likely to be a unique expression signature of each cell type. To demonstrate the utility of these expression signatures, we applied these data to the whole blood genomic profiles of several neurological diseases that we have studied previously, including acute ischemic stroke [8], migraine [9], and Tourette syndrome [10], to demonstrate that the blood genomic signatures of each of these conditions can be ascribed to certain blood cell subtypes being affected by each disease. Future studies likely could determine not only whether hematological and systemic diseases affect gene expression in specific subsets of blood cells, but also whether the diseases affect specific signaling pathways in specific subsets of cells in blood.

Results

Qualitative analysis

The numbers of "present" and "unique" genes for each cell type are listed in Table 1. Of the 54,675 genes, higher percentages of the genes are expressed (present) by lymphocytes, including B (36.6%), CD4 (36.4%), CD8 (35.3%), and NK cells (36.2%), than by monocytes/platelets (31.9%) and PMNs (24.0%). However, there are higher percentages of unique genes for PMNs (1.2%) and monocytes/platelets (0.9%) than for lymphocytes (0–0.6%) (Table 1). Among the lymphocytes, B cells have the highest number of characteristic genes, while unique transcripts for T cells are relatively scarce due to the largely similar profiles for T cell subsets $CD4^+$ and $CD8^+$ and the profiles for NK cells.

Quantitative analysis

Among 54,675 genes (probe sets) on the array, a total of 2635 are differentially expressed between the blood cell types (parametric analysis of variance (ANOVA), $p < 0.05$ with Bonferroni correction), among which 269 are significant using a Student–Newman–Keuls post hoc test. For practical reasons, we focused on the 269-probe set list since it should contain the most characteristic genes and potentially contain genomic expression markers for each cell lineage. These 269 genes/probe sets were mathematically separated into nine clusters of relatively unique expression profiles using a hierarchical algorithm [25] as demonstrated in Fig. 1. The pattern of expression of each gene in each cluster and the fold changes of the genes are shown in separate panels on the right side of Fig. 1. In general the fold changes varied as much as 10- to 100-fold. The genes in each cluster are listed in Table 2. The left side of Fig. 1 not only shows the gene expression (y axis) for different blood cell types (x axis), but also shows the gene expression of the three individuals performed for each cell type. Note that the expression levels (red—fivefold increase; bright green—fivefold decrease) for each individual are extremely reproducible between cell types and between genes. This indicates that the microarray technological variables have a minimal effect upon the expression profiles shown in Fig. 1 and indicate that our criteria for selecting genes for each cell type are stringent and not significantly affected by individual differences at least in this preliminary study.

Table 1
Results of the qualitative analysis

Cell type	Number of present probe sets	Percentage of present probe sets	Number of unique probe sets	Percentage of unique probe sets
PMN	13,139	24.0%	155	1.2%
$CD14^+$ monocyte or platelet	17,426	31.9%	152	0.9%
$CD19^+$ B cell	20,009	36.6%	120	0.6%
$CD4^+$ T cell	19,909	36.4%	24	0.1%
$CD8^+$ T cell	19,321	35.3%	8	0.0%
$CD56^+$ NK cell	19,777	36.2%	42	0.2%

A total of 54,675 probe sets that examined approximately 39,500 genes were surveyed on each array. "Present" probe sets include probe sets that have 3 present detection calls for a specific cell type regardless of the calls for other cell types. "Unique" probe sets include probe sets that have 3 present calls for a cell type and 15 absent calls for every other cell type.

Cellular origin of blood genes regulated by neurological diseases

Fig. 2 represents a melding of the data from the current study with that from our previous disease-specific studies. The genes that were most highly regulated in ischemic stroke [8], Tourette syndrome [10], and migraine [9] were selected and the cell-specific expression of each of those genes (from the present study) is shown. As demonstrated in Fig. 2, the major genes up-regulated in whole blood after stroke were expressed mainly by PMNs and monocytes/platelets. The major genes up-regulated by Tourette syndrome were mostly from NK cells and/or $CD8^+$ T cells. The major genes up-regulated by migraine were predominantly from platelets/monocytes,

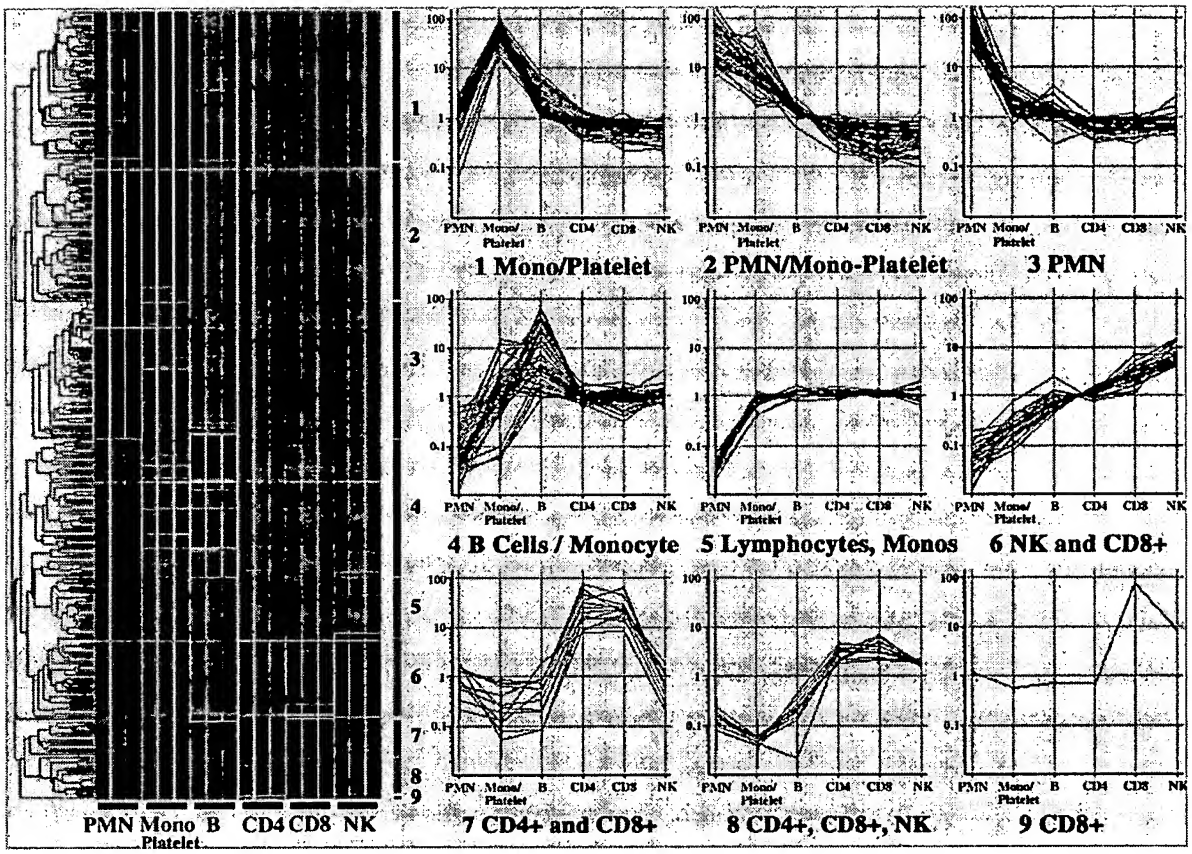


Fig. 1. A total of 269 genes that are differentially expressed between blood cell types (parametric one-way ANOVA followed by Student–Newman–Keuls post hoc test, $p < 0.05$, with Bonferroni multiple comparison correction) were subjected to a hierarchical cluster algorithm with Pearson correlation as a measure of similarity. (Left) Clusters of genes (nine clusters) with similar expression patterns are displayed from top to bottom (y axis), while cell types are displayed from left to right along the x axis. For each cell type the results of the three different individuals are shown adjacent to one another. The relative expression of each gene is color coded; red shows a fivefold increase and green shows a fivefold decrease. (Right, 1–9) Line graphs of genes segregated in the cluster analysis are shown for each of the nine clusters identified on the left. The x axis shows the cell types and the y axis shows the relative expression values (log scale) as mean −1 standard deviation (log ratio).

though there were some regulated genes from PMNs, $CD4^+$, $CD8^+$, and NK cells.

Discussion

This study surveyed the global expression profiles of six major subtypes of blood cells. These data support previous studies showing that T cells and even platelets have genes that are expressed in common, but also have genes that are fairly specific for each cell type and, perhaps more importantly, have different families of genes that tend to be expressed in a specific cell type compared to another cell type [14,18–20]. Characterization of these profiles should help elucidate the molecular and genomic basis of the development, differentiation, and function for each cell type.

Genes in cluster 1 are highly enriched in a monocyte/platelet population compared to other cell types. The recent literature shows that many genes from this cluster, such as CLU, GP1BB, PF4V1, and others (Table 2), are specifically expressed by platelets [18,20]. Cluster 2 represents genes enriched in PMNs and monocytes, while genes in cluster 3 are expressed exclusively by PMNs. Many genes in these two clusters play crucial roles in innate immunity. These include receptor molecules such as TREM1 [26], FPRL1 [27], and TLR2 [28], which are involved in microbial recognition and lead to phagocyte activation and the amplification of the inflammatory response. There are effector molecules such as MMP9 [29]; S100 proteins P, A9, and A12 [30]; and neutrophil cytosolic factors 1, 2, and 4, which participate in the neutralization of and aid clearance of microorganisms and foreign materials, and scavenger molecules such as IL1R2 [31] and TNFRSF10 [32] that help suppress excessive and harmful innate immune responses. In comparison, genes down-regulated in PMNs (cluster 5) did not provide many functional insights. The low expression of several ribosomal proteins and transcription elongation factor in this cluster may indicate a slower rate of protein translation in PMNs and is consistent with somewhat fewer RNA transcripts in this cell type (Table 1).

Several molecules expressed by B cells (cluster 4) serve important central roles in B cell development, proliferation, and differentiation, such as MS4A1 [33], BLNK [34], and BANK1 [35]. Other molecules, including immunoglobulins and HLA antigens, important for normal B cell functions, were also expressed (Table 2). While there are a few common genes between NK cells and T cells, most notably T cell receptor subunits and lymphocyte-specific tyrosine kinase (LCK)

Table 2
Results of the quantitative analysis

	Common	GenBank	Description
Cluster 1	a1/3GTP	AI972498	Clone IMAGE:4812754, mRNA
	ACRBP	AB051833	Acrosin-binding protein
	ARHGAP6	NM_001174	Rho GTPase-activating protein 6
	C21orf7	NM_020152	Chromosome 21 open reading frame 7
	CA2	M36532	Carbonic anhydrase II
	CD163	NM_004244	CD163 antigen
	CD36	NM_000072	CD36 antigen (collagen type I receptor, thrombospondin receptor)
	CD9	NM_001769	CD9 antigen (p24)
	CLEC2	NM_016509	C-type lectin-like receptor-2
	CLU	M25915	Clusterin
	CSPG2	BF590263	Chondroitin sulfate proteoglycan 2 (versican)
	CXCL5	AK026546	Chemokine (C-X-C motif) ligand 5
	CYP1B1	NM_000104	Cytochrome P450, family 1, subfamily B, polypeptide 1
	ELOVL7	AW138767	Hypothetical protein FLJ23563
	EMS1	NM_005231	EMS1 sequence (mammary tumor and squamous cell carcinoma-associated (p80/85 Src substrate)
	F13A1	NM_000129	Coagulation factor XIII, A1 polypeptide
	FSTL1	BC000055	Follistatin-like 1
	GNG11	NM_004126	Guanine nucleotide binding protein (G protein), γ11
	GP1BB	NM_000407	Glycoprotein Ib (platelet), β polypeptide
	HIST1H3H	NM_003536	Histone 1, H3h
	ITGB3	M35999	Integrin, β3 (platelet glycoprotein IIIa, antigen CD61)
	KIAA0626	NM_021647	
	MS4A6A	NM_022349	Membrane-spanning 4-domains, subfamily A, member 6A
	MYLK	AA526844	MSTP083 mRNA, complete cds
	PF4	NM_002619	Platelet factor 4 (chemokine (C-X-C motif) ligand 4)
	PF4V1	NM_002620	Platelet factor 4 variant 1
	PPBP	R64130	Proplatelet basic protein (chemokine (C-X-C motif) ligand 7)
	PRKAR2B	NM_002736	Protein kinase, cAMP-dependent, regulatory, type II, β
	PROS1	NM_000313	Protein S (α)
	PTGS1	S36219	Prostaglandin–endoperoxide synthase 1 (prostaglandin G/H synthase and cyclooxygenase)
	RIN2	AL136924	Ras and Rab interactor 2
Cluster 1	SDPR	NM_004657	Serum deprivation response (phosphatidylserine binding protein)
	SDPR	BF982174	Serum deprivation response (phosphatidylserine binding protein)
	SPARC	NM_003118	Secreted protein, acidic, cysteine-rich (osteonectin)
	THBS1	BF055462	Thrombospondin 1
	TREML1	AF534823	Triggering receptor expressed on myeloid cells-like 1
	TUBB1	NM_030773	Tubulin, β1
Cluster 2	ANXA3	M63310	Annexin A3
	APOBEC3A	U03891	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3A
	AQP9	NM_020980	Aquaporin 9
	BASP1	NM_006317	Brain abundant, membrane attached signal protein 1
	CD14	NM_000591	CD14 antigen
	CLECSF12	AF400600	C-type (calcium-dependent, carbohydrate-recognition domain) lectin, superfamily member 12
	CLECSF9	BC000715	C-type (calcium-dependent, carbohydrate-recognition domain) lectin, superfamily member 9
	CREB5	AI689210	cAMP-responsive element binding protein 5
	CSF3R	NM_000760	Colony-stimulating factor 3 receptor (granulocyte)
	DKFZP434B044	AL136861	Hypothetical protein DKFZp434B044
	DKFZp434H2111	AK026776	Hypothetical protein DKFZp434H2111
	FCGR2A	NM_021642	Fc fragment of IgG, low affinity IIa, receptor for (CD32)
	FLJ20273	NM_019027	RNA-binding protein
	FLJ23091	AL534095	Putative NF-κB activating protein 373
	FLJ23091	AA775681	Putative NF-κB activating protein 373
	FLJ23153	AA650281	Likely ortholog of mouse tumor necrosis-α-induced adipose-related protein
	FOS	BC004490	v-Fos FBJ murine osteosarcoma viral oncogene homolog
	FPR1	NM_002029	Formyl peptide receptor 1
	GALNAC4S-6ST	NM_014863	
	GPR86	NM_023914	G-protein-coupled receptor 86
	HIST2H2BE	NM_003528	Histone 2, H2be
	HSPC159	AK025603	HSPC159 protein
	IL13RA1	NM_001560	Interleukin 13 receptor, α1
	IL1RN	U65590	
	MNDA	NM_002432	Myeloid cell nuclear differentiation antigen

Table 2 (*continued*)

	Common	GenBank	Description
Cluster 2	NCF1	NM_000265	Neutrophil cytosolic factor 1 (47 kDa, chronic granulomatous disease, autosomal 1)
	NCF2	BC001606	Neutrophil cytosolic factor 2 (65 kDa, chronic granulomatous disease, autosomal 2)
	NCF4	NM_013416	Neutrophil cytosolic factor 4, 40 kDa
	NFE2	L13974	Nuclear factor (erythroid-derived 2), 45 kDa
	PADI4	NM_012387	Peptidyl arginine deiminase, type IV
	QPCT	NM_012413	Glutaminyl-peptide cyclotransferase (glutaminyl cyclase)
	RGS18	AF076642	Regulator of G-protein signaling 18
	S100A12	NM_005621	S100 calcium-binding protein A12 (calgranulin C)
	S100A9	NM_002965	S100 calcium-binding protein A9 (calgranulin B)
	SGK	NM_005627	Serum/glucocorticoid regulated kinase
	SLC22A4	NM_003059	Solute carrier family 22 (organic cation transporter), member 4
	SNCA	BG260394	Synuclein, α (non-A4 component of amyloid precursor)
	TLR2	NM_003264	Toll-like receptor 2
	TLR4	U93091	
	TLR8	AW872374	
	TM6SF1	NM_023003	Transmembrane 6 superfamily member 1
	TMG4	BF905445	Transmembrane γ-carboxyglutamic acid protein 4
	TREM1	NM_018643	Triggering receptor expressed on myeloid cells 1
Cluster 3	ABCA1	NM_005502	ATP-binding cassette, subfamily A (ABC1), member 1
	ACSL1	NM_001995	Acyl-CoA synthetase long-chain family member 1
	ADM	NM_001124	Adrenomedullin
	C4BPA	NM_000715	Complement component 4 binding protein, α
	CCR3	NM_001837	Chemokine (C-C motif) receptor 3
	CHI3L1	M80927	Chitinase 3-like 1 (cartilage glycoprotein-39)
	CKLFSF2	AA778552	Chemokine-like factor superfamily 2
	CYP4F3	NM_000896	Cytochrome P450, family 4, subfamily F, polypeptide 3
	EMR3	AF239764	EGF-like module-containing, mucin-like, hormone receptor-like 3
	G0S2	NM_015714	Putative lymphocyte G0/G1 switch gene

Table 2 (*continued*)

	Common	GenBank	Description
Cluster 3	GPR109B	NM_006018	Putative chemokine receptor
	HAL	NM_002108	Histidine ammonia-lyase
	IL1R2	U64094	Human soluble type II interleukin-1 receptor mRNA, complete cds
	IL8	NM_000584	Interleukin 8
	IL8RB	NM_001557	Interleukin 8 receptor, β
	KCNJ15	D87291	Potassium inwardly rectifying channel, subfamily J, member 15
	KCNJ2	BF111326	Potassium inwardly rectifying channel, subfamily J, member 2
	KRT23	NM_015515	Keratin 23 (histone deacetylase inducible)
	MANSC1	NM_018050	Hypothetical protein FLJ10298
	MGAM	NM_004668	Maltase–glucoamylase (α-glucosidase)
	MME	AI433463	Membrane metalloendopeptidase (neutral endopeptidase, enkephalinase, CALLA, CD10)
	MMP9	NM_004994	Matrix metalloproteinase 9 (gelatinase B, 92-kDa gelatinase, 92-kDa type IV collagenase)
	MSCP	BG251467	Mitochondrial solute carrier protein
	PBEF1	BC020691	Pre-B-cell colony enhancing factor 1
	PROK2	AF182069	Prokineticin 2
	PTGS2	NM_000963	Prostaglandin–endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)
	S100P	NM_005980	S100 calcium-binding protein P
	SEC14L1	AI017770	SEC14-like 1 (*Saccharomyces cerevisiae*)
	TNFAIP6	NM_007115	Tumor necrosis factor, α-induced protein 6
	TNFRSF10C	AF012536	Tumor necrosis factor receptor superfamily, member 10c, decoy without an intracellular domain
	VNN3	NM_018399	Vanin 3
Cluster 4	AKAP2	BG540494	Paralemmin 2
	ANXA2	NM_004039	Annexin A2
	ATP1B3	U51478	ATPase, Na^+/K^+ transporting, β3 polypeptide
	BANK1	NM_017935	B-cell scaffold protein with ankyrin repeats 1
	BLNK	NM_013314	B-cell linker
	CCDC6	AK024913	cDNA: FLJ21260 fis, clone COL01441
	CPVL	NM_031311	Carboxypeptidase, vitellogenic-like
	CXXC5	BC006428	CXXC finger 5
	DPYSL2	NM_001386	Dihydropyrimidinase-like 2
	FCRH3	BF514552	Fc receptor-like protein 3

(continued on next page)

Table 2 (*continued*)

	Common	GenBank	Description
Cluster 4	FLJ20668	AI707896	Hypothetical protein FLJ20668
	HLA-DPA1	M27487	Major histocompatibility complex, class II, DP α1
	HLA-DPB1	NM_002121	Major histocompatibility complex, class II, DP β1
	HLA-DRB3	AJ297586	Major histocompatibility complex, class II, DR β3
	ICSBP1	AI073984	Interferon consensus sequence binding protein 1
	IGLJ3	X57812	Immunoglobulin λ joining 3
	KYNU	D55639	Kynureninase (L-kynurenine hydrolase)
	LIPA	NM_000235	Lipase A, lysosomal acid, cholesterol esterase (Wolman disease)
	MGC27165	S55735	Hypothetical protein MGC27165
	MS4A1	BC002807	Membrane-spanning 4-domains, subfamily A, member 1
	MS4A7	AI301935	Membrane-spanning 4-domains, subfamily A, member 7
	NAP1L	AI763426	Napsin B pseudogene
	P2RX5	U49396	Purinergic receptor P2X, ligand-gated ion channel, 5
	POU2AF1	NM_006235	POU domain, class 2, associating factor 1
	PRDX4	NM_006406	Peroxiredoxin 4
	SPAP1	AL833361	SH2 domain-containing phosphatase anchor protein 1
	TCF4	BF592782	Transcription factor 4
	TPD52	AA524023	Tumor protein D52
	TRAF5	NM_004619	TNF receptor-associated factor 5
	TXNDC5	NM_030810	Thioredoxin domain-containing 5
	VAMP8	NM_003761	Vesicle-associated membrane protein 5 (myobrevin)
Cluster 5	C1QBP	L04636	Complement component 1, q subcomponent binding protein
	CLNS1A	AF005422	Chloride channel, nucleotide-sensitive, 1A
	DOCK10	NM_017718	
	FLJ11171	AK023183	Hypothetical protein FLJ11171
	FLJ20160	AA133311	FLJ20160 protein
	FLJ38426	BF679966	Hypothetical protein FLJ38426
	LRPPRC	AI653608	Leucine-rich PPR-motif-containing
	MGC5395	BG287862	Hypothetical protein MGC5395
	MRPL3	BC003375	Mitochondrial ribosomal protein L3
	MRPL9	BC004517	Mitochondrial ribosomal protein L9
	MRPS23	BC000242	Mitochondrial ribosomal protein S23
	NDUFB2	NM_004546	NADH dehydrogenase (ubiquinone) 1β subcomplex, 2, 8 kDa
	PLAC8	NM_016619	Placenta-specific 8

Table 2 (*continued*)

	Common	GenBank	Description
Cluster 5	PP	NM_021129	Pyrophosphatase (inorganic)
	RAFTLIN	D42043	Raft-linking protein
	RPL10A	NM_007104	Ribosomal protein L10a
	SUCLG2	AL050226	Succinate-CoA ligase, GDP-forming, β subunit
	TCERG1	NM_006706	Transcription elongation regulator 1 (CA150)
	VPS45A	AF165513	Vacuolar protein sorting 45A (yeast)
	AKR1C3	AB018580	Aldo-keto reductase family 1, member C3 (3-α hydroxysteroid dehydrogenase, type II)
	CD160	NM_007053	CD160 antigen
	CD3Z	J04132	CD3Z antigen, ζ polypeptide (TiT3 complex)
	EAT2	BC022407	SH2 domain-containing molecule EAT2
Cluster 6	GNLY	M85276	*Homo sapiens* NKG5 gene, complete cds
	GPR56	AL554008	G-protein-coupled receptor 56
	GZMB	J03189	Granzyme B (granzyme 2, cytotoxic T-lymphocyte-associated serine esterase 1)
	GZMH	M36118	Granzyme H (cathepsin G-like 2, protein h-CCPX)
	IFIX	AI827431	Interferon-inducible protein X
	KLRC3	NM_002261	Synonyms: NKG2E, NKG2-E
	KLRD1	U30610	Killer cell lectin-like receptor subfamily D, member 1
	KLRF1	NM_016523	Killer cell lectin-like receptor subfamily F, member 1
	KLRK1	AF439512	Killer cell lectin-like receptor subfamily K, member 1
	KSP37	AB021123	Ksp37 protein
	MGC61571	BE963026	Hypothetical protein MGC61571
	NKG7	NM_005601	Natural killer cell group 7 sequence
	SAMD3	AI129628	Sterile α motif domain-containing 3
	SPON2	NM_012445	Spondin 2, extracellular matrix protein
	SPUVE	NM_007173	Protease, serine, 23
	TGFBR3	NM_003243	Transforming growth factor, β receptor III (betaglycan, 300 kDa)
	TRD@	X06557 AA227879	T-cell receptor δ locus Transcribed sequences
Cluster 7	C6orf190	BC043608	Chromosome 6 open reading frame 207, mRNA (cDNA clone IMAGE: 5764019), partial cds

Table 2 (*continued*)

	Common	GenBank	Description
Cluster 7	CCR7	NM_001838	Chemokine (C-C motif) receptor 7
	CD28	NM_006139	CD28 antigen (Tp44)
	GZMK	NM_002104	Granzyme K (serine protease, granzyme 3; tryptase II)
	IL7R	NM_002185	Interleukin 7 receptor
	LEF1	AF288571	Lymphoid enhancer-binding factor 1
	MAL	NM_002371	Mal, T-cell differentiation protein
	NELL2	NM_006159	NEL-like 2 (chicken)
	RGS1	S59049	Regulator of G-protein signaling 1
	TRIM	AJ240085	T-cell receptor-interacting molecule
		M12959	Human mRNA for T-cell receptor α chain
Cluster 8	BCL11B	AA918317	B-cell CLL/lymphoma 11B (zinc finger protein)
	CD2	NM_001767	CD2 antigen (p50), sheep red blood cell receptor
	H963	NM_013308	Platelet-activating receptor homolog
	LCK	NM_005356	Lymphocyte-specific protein tyrosine kinase
	LOC283666	AW006185	Hypothetical protein LOC283666, mRNA (cDNA clone IMAGE: 4415549), partial cds
	TRGV9	M30894	T-cell receptor γ locus
	TRGV9	M16768	T-cell receptor (V-J-C) precursor;
		BC040965	Clone IMAGE:5747561, mRNA
		AF043179	*H. sapiens* T-cell receptor β chain (TCRBV13S1–TCRBJ2S1) mRNA, complete cds
		AL559122	T-cell receptor β chain BV20S1 BJ1-5 BC1 mRNA, complete cds
		M15564	T-cell receptor precursor; human T-cell receptor rearranged β-chain V-region (V-D-J) mRNA
Cluster 9	CD8A	AW006735	CD8 antigen, α polypeptide (p32)

A total of 269 genes that are differentially expressed between blood cell types (parametric one-way ANOVA followed by Student–Newman–Keuls post hoc test, $p < 0.05$, with Bonferroni multiple correction) were segregated into nine clusters using a hierarchical cluster algorithm. Cluster 1, specifically expressed by monocytes/platelets. Cluster 2, highly expressed by PMNs and, to a lesser degree, monocytes/platelets. Cluster 3, specifically expressed by PMNs. Cluster 4, highly expressed by B cells and, to a lesser degree, monocytes/platelets. Cluster 5, expressed by all cell types except PMNs. Cluster 6, highly expressed by NK cells and, to a lesser degree, $CD8^+$ T cells. Cluster 7, highly expressed by $CD4^+$ and $CD8^+$ T cells. Cluster 8, highly expressed by $CD4^+$ and $CD8^+$ T cells and NK cells. Cluster 9, highly expressed by $CD8^+$ T cells and NK cells. Note that multiple probe sets for the same gene are presented only once.

(cluster 8), NK cells can be mostly distinguished by the higher expression of genes including NKG5, NKG7, NKG2E, KLRD, and KLRF1 (cluster 6). Interestingly, genes originally cloned from cytotoxic T cells such as GZMB and GZMH (granzyme B and H) have a higher expression in NK cells but are also present in $CD8^+$ T cells (Figs. 1 and 2; Table 2). The overall genomic patterns of $CD4^+$ and $CD8^+$ T cells are very similar and can be distinguished from other cell types by well-known T cell markers such as CD28 [36], IL7R [37], the chemokine receptor CCR7 [38], and others (cluster 7). Moreover, a novel gene, NELL2 [39], that is abundant in neural tissues is also highly expressed by both $CD4^+$ and $CD8^+$ T cells. Not surprisingly, CD8 antigen distinguishes $CD4^+$ and $CD8^+$ T cells (cluster 9). In a recent study of T cells purified from buffy coats from healthy donors, approximately 50% of the genes in clusters 7 and 8 in this study were present (6/11 and 5/11, respectively) [14], suggesting that these sets of genes are consistent even with different purification methodologies and different individuals.

Several limitations of the current study need to be emphasized. The purification of most of the cell types was less than desirable. However, it was for this reason that we identified gene clusters that are quite different from each other and then examined how these gene clusters related to the imperfectly separated cell types in blood (Fig. 1). Using this approach we can say that the genes or gene clusters that are expressed only for a given cell type are in fact expressed by that cell type. Genes that appear to be expressed by two or more cell types might be shared between those cell types or could be expressed by only one cell type but could not be differentiated because of imprecise separation of cells. In addition, a given cell type is likely to have heterogeneous subtypes, and thus there are likely to be different gene expression signatures for these subtypes. A good example is $CD4^+$ cells, among which are Th1 and Th2 $CD4^+$ cells, which have very different functions in the immune system [40] and which were not differentiated in this study. Moreover, the increased expression by any given cell type could be due to a number of different factors including the numbers of cells. This is not likely to be a major factor since the changes in gene expression shown usually vary from 5- to 100-fold, and it is unlikely that the numbers of these types of cells vary this much.

Other limitations of the results include the negative selection method of isolating the cells. Even though the selected cells do not undergo antibody binding used for positive selection, the negative selection likely activates complement that could affect the cells being surveyed. In addition, because the negative selection method does not produce pure populations of cells, it is possible that the contaminating cells activate the primary cells being isolated. Therefore, any given profile for each cell type could represent some degree of activation of the cells. Since the data at least for the three individuals examined look fairly consistent, if there is nonspecific activation of cells with negative selection at least it is fairly consistent from one experiment to the next.

An additional limitation of the current study was the inability to distinguish gene expression by platelets from gene expression by monocytes because of the negative selection methods used.

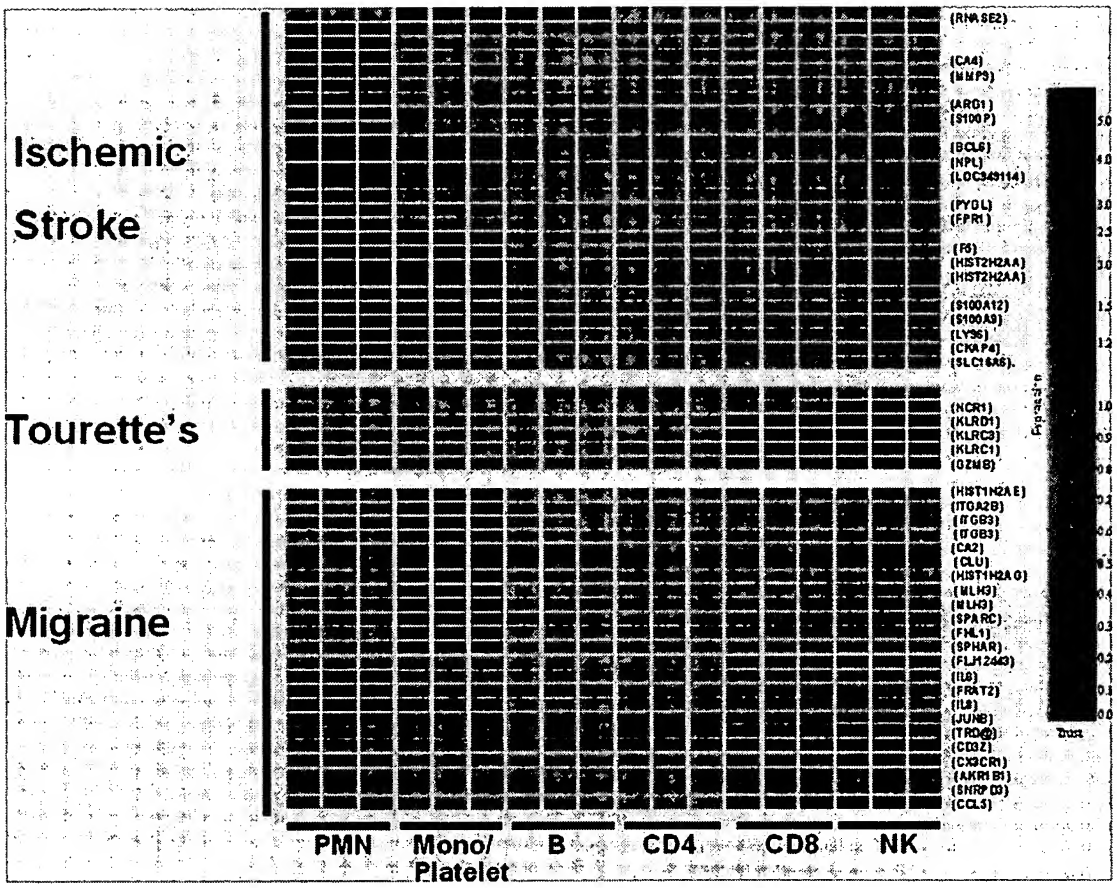


Fig. 2. Cellular origins of genes that have previously been shown to be regulated by ischemic stroke [8], Tourette syndrome [10], and migraine [9]. These genes were identified from RNA isolated from whole blood by comparing to healthy controls and patients with other diseases. Genes regulated by each disease are listed from top to bottom (*y* axis) and their relative expression in each blood cell type is color coded and displayed from left to right (*x* axis). The list of displayed genes from the top down for each disorder follows and the genes are numbered and named according to the Affymetrix probe sets: ischemic stroke, 206111_at (RNASE2), 239893_at, 234632_x_at, 206209_s_at (CA4), 203936_s_at (MMP9), 244218_at, 206177_s_at (ARG1), 204351_at (S100P), 232595_at, 228758_at (BCL6), 240440_at (NPL), 225899_x_at (LOC349114), 227129_x_at, 232958_at (PYGL), 205118_at (FPR1), 228642_at, 231029_at, 204714_s_at (F5), 214290_s_at (HIST2H2AA), 218280_x_at (HIST2H2AA), 222303_at, 205863_at (S100A12), 203535_at (S100A9), 206584_at (LY96), 200999_s_at (CKAP4), 207038_at (SLC16A6); Tourette syndrome, 212775_at (KIAA0657), 207860_at (NCR1), 207796_x_at (KLRD1), 207723_s_at (KLRC3), 206785_s_at (KLRC1), and 210164_at (GZMB); Migraine, 214469_at (HIST1H2AE), 206494_s_at (ITGA2B), 204627_s_at (ITGB3), 216261_at (ITGB3), 209301_at (CA2), 208791_at (CLU), 207156_at (HIST1H2AG), 217216_x_at (MLH3), 204838_s_at (MLH3), 200665_s_at (SPARC), 210299_s_at (FHL1), 206272_at (SPHAR), 201818_at (FLJ12443), 211506_s_at (IL8), 209864_at (FRAT2), 202859_x_at (IL8), 201473_at (JUNB), 217143_s_at (TRD@), 210031_at (CD3Z), 205898_at (CX3CR1), 201272_at (AKR1B1), 202567_at (SNRPD3) 204655_at (CCL5).

However, gene expression profiles for platelets have been published [18,20]. It is notable that platelets lack a nucleus and nuclear DNA, and hence the mRNA found in platelets is derived from the megakaryocytes that form the platelets. A similar situation exists for red blood cells, which contain mRNA but have no nucleus or nuclear DNA. The importance of differentiating platelet from monocyte RNAs is emphasized by the data in Fig. 2. Even though stroke and migraine appear to express genes in both platelets and monocytes, careful examination of the gene lists for both shows that in stroke the genes induced are mainly in monocytes, and in migraine the genes are induced mainly in platelets (see below). It is certainly possible that a genetic disease like migraine could be associated with changes in gene expression in platelets and not other cellular elements in the blood. This is emphasized by a number of clinical studies in migraine as mentioned next.

Finally, it is important to reemphasize that the data shown in this study are for expression in three normal, healthy individuals, and the data for the three diseases were obtained by comparing the published disease-regulated genes to the genes expressed by different cell types in the healthy persons in this study. The expression profiles for the individual cell types shown here can be viewed only as being preliminary, since only three individuals were studied. Large numbers of individuals may be necessary to derive reliable gene profiles for individual cell types in blood, since age, gender, race, genetic background, lifestyle, diet, concurrent diseases and medications, and many other factors are likely to influence cell type-related gene expression. The current study does, however, emphasize the need for future studies to isolate these cell types in individuals with each disease and replicate the gene expression profiles for each cell type in each

individual with each disease. An unexpected benefit to isolating individual cell types was that the fold differences of gene expression is 10- to 100-fold, compared to fold changes in whole blood, at least in neurological diseases, on the order of 2- to 3-fold [6–10,13–17,41]. The ability to detect high fold changes in these cell-specific data could be due in part to removal of high-abundance RNAs that populate whole blood and in part to using an approach that highlights cell differences rather than averaging them.

The data shown suggest that the cell type-specific data could serve as a powerful guide to understanding the relative contribution of each cell type to the overall gene expression profile caused by various diseases. The data show that different neurological diseases affect gene expression through distinct blood cell populations. Tourette syndrome, a neuropsychiatric tic disorder that has been proposed to be caused by an autoimmune response to streptococcus at least in some patients [42,43], appears to be associated with changes in gene expression in NK cells and/or $CD8^+$ cells based upon the results of this study and our previous findings [10]. In contrast, the blood genomic response following acute ischemic stroke was predominantly from PMNs based upon the current findings and those of our recent ischemic stroke study [8]. This agrees with human and animal studies showing that polymorphonuclear cells are the major cell type that initially infiltrate areas of cerebral infarction following stroke and appear to play a major role in pathogenesis [44–46]. Importantly, infection and predisposition to inflammation may be risk factors for stroke [47]. In comparison, the blood genomic pattern for migraine headache is more heterogeneous, with several blood cell types affected, including platelets and monocytes. Platelets and abnormalities of serotonin in platelets have long been implicated in the pathogenesis of migraine and its genetic basis and may contribute to an increased risk of stroke [48–51]. These data suggest that although neuroinflammation plays a crucial role in the pathological process of both cerebral ischemia [52] and migraine [53], the involved blood cells and genes are distinct. This further underscores the importance of identifying the blood cell types associated with a given disorder, to formulate mechanistic hypotheses and accurately characterize surrogate blood markers for diagnostic, prognostic, and treatment purposes.

Materials and methods

Separation of blood cell subtypes

Blood was drawn from three healthy donors for all cell subsets. The donors were healthy, ages 32, 33, and 59; two were male; and none had any concurrent infection or major medical illnesses. Subsets of mononuclear cells including $CD4^+$ T cells, $CD8^+$ T cells, $CD19^+$ B cells, $CD56^+$ NK cells, and $CD14^+$ monocytes were enriched using the RosetteSep negative selection method (StemCell Technologies, BC, Canada). This method was chosen as it provided a short processing time and the desired cell types remain largely undisturbed. This method thus helped to minimize potential alterations in gene expression due to excessive handling during the cell isolation and to avoid possible effects of positive selection of cells on gene expression [21,22]. For each cell subtype, 8 ml whole blood was drawn into Vacutainer CPT cell preparation tubes that contained sodium citrate and Ficoll (Becton–Dickinson, NJ, USA). Four hundred microliters of RosetteSep antibody cocktail was added to each tube and incubated for 20 min at room temperature. The antibody cocktail cross-links unwanted cells in whole blood to red blood cells, forming immunorosettes with increased density. After incubation, the blood collected in the CPT tubes was centrifuged for 20 min at 1800 g to precipitate the unwanted cells and precipitate free RBCs. The desired cells, not labeled with antibody, were collected from the plasma:Ficoll interface and washed once with PBS and the RNA was isolated using Trizol reagent. The purity of cells separated with this protocol was 90 ± 5% for $CD4^+$ T cells, 76 ± 8% for $CD8^+$ T cells, 81 ± 8% for $CD19^+$ B cells, 74 ± 10% for $CD56^+$ NK cells, and 69 ± 12% for $CD14^+$ monocytes (mean ± SD) according to confirmatory flow cytometry data provided by the manufacturer.

It is noted that platelet contamination has been reported for $CD14^+$ monocytes using the RosetteSep enrichment method. Therefore, the "monocyte" gene expression pattern obtained in this study may come from both monocytes and platelets and is therefore referred to as "mono/platelet" in the text. However, importantly, platelet contamination is not seen in lymphocyte enrichments as the lymphocyte enrichment cocktails contain anti-CD36, which removes platelets along with monocytes.

For the separation of PMNs, 6 ml sodium citrate–anti-coagulated venous blood was carefully layered on 3 ml leukocyte separation media (Histopaque-1119 and Histopaque-1077; Sigma–Aldrich, St. Louis, MO, USA). After a 700*g* centrifugation for 30 min at room temperature, the "mononuclear" layer together with fluid within 0.5 cm of this layer was aspirated and discarded, while the "granulocyte" layer was transferred to a new tube and washed once with phosphate-buffered saline. The purity of the PMNs obtained was ~98% based on microscopic examination after Wright–Giemsa staining.

RNA isolation

The collected pellets of each cell subtype were immediately homogenized with Trizol reagent (Invitrogen, Carlsbad, CA, USA). The aqueous phase containing RNA was separated by centrifugation after adding chloroform. RNA was recovered by precipitation with isopropyl alcohol and washed with 75% ethanol. RNA samples were quantified using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) and the quality was assessed using an Agilent 2100 bioanalyzer (Agilent, Palo Alto, CA, USA). We processed RNA only when the A_{260}/A_{280} absorbance ratio of the cleaned RNA exceeded 2.0 and 28S/18S ratio equaled or exceeded 1.8 for these microarray studies.

Microarray processing

For each blood cell subtype ($CD4^+$ T cells, $CD8^+$ T cells, $CD19^+$ B cells, $CD56^+$ NK cells, $CD14^+$ monocytes, and PMNs), 25 ng total RNA was labeled using a two-cycle target labeling protocol and hybridized to arrays. A total of 36 (6 arrays for each cell type, for three subjects, with technical duplicates done for each) Affymetrix Human 2.0 Plus arrays were used (Affymetrix, Santa Clara, CA, USA) in this study, with each array containing 54,675 probe sets referred to as such or as "genes" in the text. Six arrays for each cell type were used for each of the three subjects (n = 18 arrays), and technical duplicates of these were then processed (n = 36 arrays total). The probe sets were on one chip that surveyed 47,000 transcripts from ~39,500 potential human genes (Affymetrix Technical Manual). Sample labeling, hybridization to chips, and image scanning were performed according to the Affymetrix *Expression Analysis Technical Manual.*

Data analysis

After the arrays were scanned, the raw expression values (probe level data) for each gene were saved in Affymetrix.cel and Affymetrix.dat files. The probe level data were then collated using GC-Robust Multi-array Average (GCRMA-EB) software (http://www.bioconductor.org/). This involved nonlinear background reduction, quantile normalization, and summarization by median polishing [23,24]. The technical replicates were averaged,

and therefore the analyses were performed on 18 independent data sets (three individuals and six different cell types for each individual). We used both qualitative and quantitative methods to assess the expression signatures of each cell type.

Qualitative method

This method was used to determine genes that are uniquely expressed by each cell type. Affymetrix GCOS software was used to generate "detection calls," i.e., "present," "absent," or "marginal," for each gene on each array based on Wilcoxon's signed rank test (http://www.affymetrix.com/support/technical/technotes/statistical_reference_guide.pdf). Briefly, the present genes represent transcripts that can be clearly detected by arrays, while the absent calls are generally given to transcripts that are below the detection threshold and cannot be reliably distinguished from noise. We determined the number of *genes that are present in each cell type* (with 3 present calls in three arrays regardless of calls from other cell types) and the number of *genes that are present only in each specific cell type* (with 3 present calls in three arrays and 15 absent calls in all other cell types). This method tends to eliminate genes that are marginally associated with a given cell type. As noted above, the technical replicates were averaged together for this qualitative analysis and the quantitative analysis that follows.

Quantitative methods

The probe level data from .cel files were first collated using Robust Multi-array Average (RMA) software (http://www.bioconductor.org/). After nonlinear background reduction and quantile normalization and summarization, probe level data were generated by median polishing each gene to produce a quantitative expression assessment in each cell sample [23,24]. Genes that were differentially expressed among cell subtypes were identified using a parametric one-way ANOVA followed by the Student–Newman–Keuls post hoc test with subtype as the variable (GENESPRING 7 software; Silicon Genetics, Redwood City, CA, USA). The type I error probability value (p value) was corrected with the Bonferroni method for multiple test comparisons and $p < 0.05$ was considered significant. The identified genes were subjected to a hierarchical cluster analysis using Pearson correlation as the similarity measure (Fig. 1).

Cellular origin of genes regulated by ischemic stroke, migraine, and Tourette syndrome

The genes up-regulated by ischemic stroke, migraine, and Tourette syndrome were identified in three previous studies by comparing these conditions to healthy controls and patients with other neurological disorders. RNA isolated from whole blood was used for all of these previous studies. The ischemic stroke study was performed using U133 2.0 Plus arrays [8], while the Tourette and migraine studies were carried out using human U95Av 2 arrays [9,10]. For the ischemic stroke study 45 stroke samples were compared to 14 healthy control samples [8]. For the migraine headache study 22 patients with migraine were compared to 56 control samples from patients with a variety of other disorders [9]. For the Tourette syndrome study, 16 patients with Tourette syndrome who had at least one first degree relative with Tourette syndrome were compared to 113 samples from control patients with a variety of other disorders including epilepsy and headache and healthy controls [10]. With the current subtype-specific expression data, we attempted to address the cell types that are affected by each disease. To make a direct comparison, genes from Affymetrix U95Av 2 arrays were matched to genes on the Affymetrix U133 2.0 Plus arrays through the Netaffx database (www.affymetrix.com) and their expression across different blood cell types was examined and visualized (Fig. 2).

Acknowledgments

This study was supported by National Institutes of Health Grants NS043252, NS 028167, NS042774, and NS044283 (F. R.S.). This work was partially supported by the MIND Institute Genomics Core and Grant PO1 ES 11269 from the NIEHS and U.S. EPA.

References

[1] T.R. Golub, D.K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J.P. Mesirov, H. Coller, M.L. Loh, J.R. Downing, M.A. Caligiuri, C.D. Bloomfield, E.S. Lander, Molecular classification of cancer: class discovery and class prediction by gene expression monitoring, Science 286 (1999) 531–537.

[2] A.A. Alizadeh, M.B. Eisen, R.E. Davis, C. Ma, I.S. Lossos, A. Rosenwald, J.C. Boldrick, H. Sabet, T. Tran, X. Yu, J.I. Powell, L. Yang, G.E. Marti, T. Moore, J. Hudson Jr., L. Lu, D.B. Lewis, R. Tibshirani, G. Sherlock, W.C. Chan, T.C. Greiner, D.D. Weisenburger, J.O. Armitage, R. Warnke, L.M. Staudt, et al., Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling, Nature 403 (2000) 503–511.

[3] L. Bennett, A.K. Palucka, E. Arce, V. Cantrell, J. Borvak, J. Banchereau, V. Pascual, Interferon and granulopoiesis signatures in systemic lupus erythematosus blood, J. Exp. Med. 197 (2003) 711–723.

[4] R.A. Heller, M. Schena, A. Chai, D. Shalon, T. Bedilion, J. Gilmore, D.E. Woolley, R.W. Davis, Discovery and analysis of inflammatory disease-related genes using cDNA microarrays, Proc. Natl. Acad. Sci. USA 94 (1997) 2150–2155.

[5] Y. Tang, A. Lu, B.J. Aronow, F.R. Sharp, Blood genomic responses differ after stroke, seizures, hypoglycemia, and hypoxia: blood genomic fingerprints of disease, Ann. Neurol. 50 (2001) 699–707.

[6] A.R. Whitney, M. Diehn, S.J. Popper, A.A. Alizadeh, J.C. Boldrick, D.A. Relman, P.O. Brown, Individuality and variation in gene expression patterns in human blood, Proc. Natl. Acad. Sci. USA 100 (2003) 1896–2191.

[7] Y. Tang, A. Lu, R. Ran, B.J. Aronow, E.K. Schorry, R.J. Hopkin, D.L. Gilbert, T.A. Glauser, A.D. Hershey, N.W. Richtand, M. Privitera, A. Dalvi, A. Sahay, J.P. Szaflarski, D.M. Ficker, N. Ratner, F.R. Sharp, Human blood genomics: distinct profiles for gender, age and neurofibromatosis type 1, Brain Res. Mol. Brain Res. 132 (2004) 155–167.

[8] Y. Tang, H. Xu, X.L. Du, W. Walker, A. Lu, R. Ran, J.P. Gregg, M. Reilly, A. Pancioli, J.C. Khoury, L.R. Sauerbeck, J.A. Carrozzella, J. Spilker, J.F. Clark, K.R. Wagner, E.C. Jauch, P. Verro, D.J. Chang, J.P. Broderick, F.R. Sharp, Gene expression in blood changes rapidly in neutrophils and monocytes following ischemic stroke in humans: a microarray study, J. Cereb. Blood Flow Metab. (in press).

[9] A.D. Hershey, Y. Tang, S.W. Powers, M.A. Kabbouche, D.L. Gilbert, T.A. Glauser, F.R. Sharp, Genomic abnormalities in patients with migraine and chronic migraine: preliminary blood gene expression suggests platelet abnormalities, Headache 44 (2004) 994–1004.

[10] Y. Tang, D.L. Gilbert, T.A. Glauser, A.D. Hershey, F.R. Sharp, Blood gene expression profiling of neurologic diseases: a pilot microarray study, Arch. Neurol. 62 (2005) 210–215.

[11] N.C. Twine, J.A. Stover, B. Marshall, G. Dukart, M. Hidalgo, W. Stadler, T. Logan, J. Dutcher, G. Hudes, A.J. Domer, D.K. Slonim, W.L. Trepicchio, M.E. Burczynski, Disease-associated expression profiles in peripheral blood mononuclear cells from patients with advanced renal cell carcinoma, Cancer Res. 63 (2003) 6069–6075.

[12] A.H. Iglesias, S. Camelo, D. Hwang, R. Villanueva, G. Stephanopoulos, F. Dangond, Microarray detection of E2F pathway activation and other targets in multiple sclerosis peripheral blood mononuclear cells, J. Neuroimmunol. 150 (2004) 163–177.

[13] M.S. Forrest, Q. Lan, A.E. Hubbard, L. Zhang, R. Vermeulen, X. Zhao, G. Li, Y.Y. Wu, M. Shen, S. Yin, S.J. Chanock, N. Rothman, M.T. Smith, Discovery of novel biomarkers by microarray analysis of peripheral blood mononuclear cell gene expression in benzene-exposed workers, Environ. Health Perspect. 113 (2005) 801–807.

[14] J.P. Cobb, M.N. Mindrinos, C. Miller-Graziano, S.E. Calvano, H.V. Baker, W. Xiao, K. Laudanski, B.H. Brownstein, C.M. Elson, D.L. Hayden, D.N. Herndon, S.F. Lowry, R.V. Maier, D.A. Schoenfeld, L.L. Moldawer, R.W. Davis, R.G. Tompkins, P. Bankey, T. Billiar, D. Camp, I. Chaudry, B. Freeman, R. Gamelli, N. Gibran, B. Harbrecht, W. Heagy, D. Heimbach, J.

Horton, J. Hunt, J. Lederer, J. Mannick, B. McKinley, J. Minei, E. Moore, F. Moore, R. Munford, A. Nathens, G. O'Keefe, G. Purdue, L. Rahme, D. Remick, M. Sailors, M. Shapiro, G. Silver, R. Smith, G. Stephanopoulos, G. Stormo, M. Toner, S. Warren, M. West, S. Wolfe, V. Young, Application of genome-wide expression analysis to human health and disease, Proc. Natl. Acad. Sci. USA 102 (2005) 4801–4806.

[15] Y. Tang, M.B. Schapiro, D.N. Franz, B.J. Patterson, F.J. Hickey, E.K. Schorry, R.J. Hopkin, M. Wylie, T. Narayan, T.A. Glauser, D.L. Gilbert, A. D. Hershey, F.R. Sharp, Blood expression profiles for tuberous sclerosis complex 2, neurofibromatosis type 1, and Down's syndrome, Ann. Neurol. 56 (2004) 808–814.

[16] F. Borovecki, L. Lovrecic, J. Zhou, H. Jeong, F. Then, H.D. Rosas, S. M. Hersch, P. Hogarth, B. Bouzou, R.V. Jensen, D. Krainc, Genome-wide expression profiling of human blood reveals biomarkers for Huntington's disease, Proc. Natl. Acad. Sci. USA 102 (2005) 11023–11028.

[17] D.F. Moore, H. Li, N. Jeffries, V. Wright, R.A. Cooper Jr., A. Elkahloun, M.P. Gelderman, E. Zudaire, G. Blevins, H. Yu, E. Goldin, A.E. Baird, Using peripheral blood mononuclear cells to determine a gene expression profile of acute ischemic stroke: a pilot investigation, Circulation 111 (2005) 212–221.

[18] P. Bugert, A. Dugrillon, A. Gunaydin, H. Eichler, H. Kluter, Messenger RNA profiling of human platelets by microarray hybridization, Thromb. Haemostasis 90 (2003) 738–748.

[19] F.B. Stentz, A.E. Kitabchi, Transcriptome and proteome expression in activated human CD4 and CD8 T-lymphocytes, Biochem. Biophys. Res. Commun. 324 (2004) 692–696.

[20] D.V. Gnatenko, J.J. Dunn, S.R. McCorkle, D. Weissmann, P.L. Perrotta, W.F. Bahou, Transcript profiling of human platelets using microarray and serial analysis of gene expression, Blood 101 (2003) 2285–2293.

[21] E. Elkord, P.E. Williams, H. Kynaston, A.W. Rowbottom, Human monocyte isolation methods influence cytokine production from in vitro generated dendritic cells, Immunology 114 (2005) 204–212.

[22] S. Matsuzaki, K. Shinozaki, N. Kobayashi, K. Agematsu, Polarization of Th1/Th2 in human CD4 T cells separated by CD62L: analysis by transcription factors, Allergy 60 (2005) 780–787.

[23] B.M. Bolstad, R.A. Irizarry, M. Astrand, T.P. Speed, A comparison of normalization methods for high density oligonucleotide array data based on variance and bias, Bioinformatics 19 (2003) 185–193.

[24] R.A. Irizarry, B.M. Bolstad, F. Collin, L.M. Cope, B. Hobbs, T.P. Speed, M. Astrand, Summaries of Affymetrix Genechip probe level data: a comparison of normalization methods for high density oligonucleotide array data based on variance and bias, Nucleic Acids Res. 31 (2003) e15.

[25] M.B. Eisen, P.T. Spellman, P.O. Brown, D. Botstein, Cluster analysis and display of genome-wide expression patterns, Proc. Natl. Acad. Sci. USA 95 (1998) 14863–14868.

[26] S. Knapp, S. Gibot, A. de Vos, H.H. Versteeg, M. Colonna, T. van der Poll, Cutting edge: expression patterns of surface and soluble triggering receptor expressed on myeloid cells-1 in human endotoxemia, J. Immunol. 173 (2004) 7131–7134.

[27] Y. Le, P.M. Murphy, J.M. Wang, Formyl-peptide receptors revisited, Trends Immunol. 23 (2002) 541–548.

[28] A.O. Aliprantis, R.B. Yang, M.R. Mark, S. Suggett, B. Devaux, J.D. Radolf, G.R. Klimpel, P. Godowski, A. Zychlinsky, Cell activation and apoptosis by bacterial lipoproteins through Toll-like receptor-2, Science 285 (1999) 736–739.

[29] G. Opdenakker, P.E. Van den Steen, B. Dubois, I. Nelissen, E. Van Coillie, S. Masure, P. Proost, J. Van Damme, Gelatinase B functions as regulator and effector in leukocyte biology, J. Leukocyte Biol. 69 (2001) 851–859.

[30] J. Roth, T. Vogl, C. Sorg, C. Sunderkotter, Phagocyte-specific S100 proteins: a novel group of proinflammatory molecules, Trends Immunol. 24 (2003) 155–158.

[31] E. Bourke, A. Cassetti, A. Villa, E. Fadlon, F. Colotta, A. Mantovani, IL-1 beta scavenging by the type II IL-1 decoy receptor in human neutrophils, J. Immunol. 170 (2003) 5999–6005.

[32] J.P. Sheridan, S.A. Marsters, R.M. Pitti, A. Gurney, M. Skubatch, D. Baldwin, L. Ramakrishnan, C.L. Gray, K. Baker, W.I. Wood, A.D. Goddard, P. Godowski, A. Ashkenazi, Control of TRAIL-induced apoptosis by a family of signaling and decoy receptors, Science 277 (1997) 818–821.

[33] T.F. Tedder, M. Streuli, S.F. Schlossman, H. Saito, Isolation and structure of a cDNA encoding the B1 (CD20) cell-surface antigen of human B lymphocytes, Proc. Natl. Acad. Sci. USA 85 (1988) 208–212.

[34] C. Fu, C.W. Turck, T. Kurosaki, A.C. Chan, BLNK: a central linker protein in B cell activation, Immunity 9 (1998) 93–103.

[35] K. Yokoyama, I.H. Su, T. Tezuka, T. Yasuda, K. Mikoshiba, A. Tarakhovsky, T. Yamamoto, Bank regulates Bcr-induced calcium mobilization by promoting tyrosine phosphorylation of IP(3) receptor, EMBO J. 21 (2002) 83–92.

[36] G. Magistrelli, P. Jeannin, G. Elson, J.F. Gauchat, T.N. Nguyen, J.Y. Bonnefoy, Y. Delneste, Identification of three alternatively spliced variants of human CD28 mRNA, Biochem. Biophys. Res. Commun. 259 (1999) 34–37.

[37] S.M. Kaech, J.T. Tan, E.J. Wherry, B.T. Konieczny, C.D. Surh, R. Ahmed, Selective expression of the interleukin 7 receptor identifies effector CD8 T cells that give rise to long-lived memory cells, Nat. Immunol. 4 (2003) 1191–1198.

[38] F. Sallusto, D. Lenig, R. Forster, M. Lipp, A. Lanzavecchia, Two subsets of memory T lymphocytes with distinct homing potentials and effector functions, Nature 401 (1999) 708–712.

[39] B.R. Nelson, S. Matsuhashi, F. Lefcort, Restricted neural epidermal growth factor-like like 2 (nell2) expression during muscle and neuronal differentiation, Gene Expression Patterns 2 (2002) 7–15.

[40] J. Parkin, B. Cohen, An overview of the immune system, Lancet 357 (2001) 1777–1789.

[41] Y. Tang, T.A. Glauser, D.L. Gilbert, A.D. Hershey, M.D. Privitera, D.M. Ficker, J.P. Szaflarski, F.R. Sharp, Valproic acid blood genomic expression patterns in children with epilepsy—A pilot study, Acta Neurol. Scand. 109 (2004) 159–168.

[42] R.R. Trifiletti, A.M. Packard, Immune mechanisms in pediatric neuropsychiatric disorders: Tourette's syndrome, OCD, and PANDAS, Child Adolesc. Psychiatr. Clin. North Am. 8 (1999) 767–775.

[43] S.E. Swedo, H.L. Leonard, B.B. Mittleman, A.J. Allen, J.L. Rapoport, S.P. Dow, M.E. Kanter, F. Chapman, J. Zabriskie, Identification of children with pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections by a marker associated with rheumatic fever, Am. J. Psychiatry 154 (1997) 110–112.

[44] K. Kariko, D. Weissman, F.A. Welsh, Inhibition of Toll-like receptor and cytokine signaling—A unifying theme in ischemic tolerance, J. Cereb. Blood Flow Metab. 24 (2004) 1288–1304.

[45] M. Ishikawa, J.H. Zhang, A. Nanda, D.N. Granger, Inflammatory responses to ischemia and reperfusion in the cerebral microcirculation, Front. Biosci. 9 (2004) 1339–1347.

[46] A. Chamorro, Role of inflammation in stroke and atherothrombosis, Cerebrovasc. Dis. 17 (Suppl. 3) (2004) 1–5.

[47] P.J. Lindsberg, A.J. Grau, Inflammation and infections as risk factors for ischemic stroke, Stroke 34 (2003) 2518–2532.

[48] J.A. Zeller, K. Frahm, R. Baron, R. Stingele, G. Deuschl, Platelet–leukocyte interaction and platelet activation in migraine: a link to ischemic stroke? J. Neurol. Neurosurg. Psychiatry 75 (2004) 984–987.

[49] M. Estevez, K.L. Gardner, Update on the genetics of migraine, Hum. Genet. 114 (2004) 225–235.

[50] M. Etminan, B. Takkouche, F.C. Isorna, A. Samii, Risk of ischaemic stroke in people with migraine: systematic review and meta-analysis of observational studies, BMJ 330 (2005) 63.

[51] M. Wessman, M.A. Kaunisto, M. Kallela, A. Palotie, The molecular genetics of migraine, Ann. Med. 36 (2004) 462–473.

[52] C. Iadecola, M. Alexander, Cerebral ischemia and inflammation, Curr. Opin. Neurol. 14 (2001) 89–94.

[53] M.A. Moskowitz, Neurogenic inflammation in the pathophysiology and treatment of migraine, Neurology 43 (Suppl. 3) (1993) S16–S20.